



日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 4 年 1 月 7 日
Date of Application:

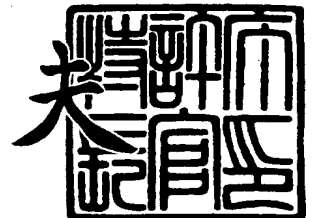
出 願 番 号 特 願 2 0 0 4 - 0 0 2 0 3 7
Application Number:
[ST. 10/C]: [J P 2 0 0 4 - 0 0 2 0 3 7]

出 願 人 株式会社日立製作所
Applicant(s):

2 0 0 4 年 3 月 2 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫





【書類名】 特許願
【整理番号】 340301258
【提出日】 平成16年 1月 7日
【あて先】 特許庁長官殿
【国際特許分類】 G06F 11/16
【発明者】
 【住所又は居所】 神奈川県小田原市中里 3 2 2 番 2 号 株式会社日立製作所 S A N
 ソリューション事業部内
 【氏名】 山下 真一郎
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社日立製作所
【代理人】
 【識別番号】 110000176
 【氏名又は名称】 一色国際特許業務法人
 【代表者】 一色 健輔
【手数料の表示】
 【予納台帳番号】 211868
 【納付金額】 21,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項 1】**

データを記憶する第 1 の記憶ボリュームを有する第 1 のストレージ装置と、
前記第 1 のストレージ装置と通信可能に接続され、データを記憶する第 2 の記憶ボリュームを有する第 2 のストレージ装置と、
を備え、
前記第 1 のストレージ装置は、
前記第 1 の記憶ボリュームにデータが書き込まれると、前記データの複製を前記第 2 のストレージ装置に送信する複製データ送信部
を備え、
前記第 2 のストレージ装置は、
前記複製データ送信部により送信される前記データの複製を受信し、前記データの複製を前記第 2 の記憶ボリュームに書き込む複製データ受信部
を備え、
前記第 1 のストレージ装置は、
第 1 のハートビートメッセージを前記第 1 の記憶ボリュームに所定時間以内の間隔で繰り返し書き込むディスクハートビート書き込み部
を備え、
前記第 2 のストレージ装置は、
前記複製データ受信部により前記第 2 の記憶ボリュームへ書き込まれる前記第 1 のハートビートメッセージの複製の検知を行うディスクハートビート検知部
を備えることを特徴とするストレージシステム。

【請求項 2】

前記第 1 のストレージ装置には第 1 の情報処理装置が通信可能に接続され、
前記第 2 のストレージ装置には第 2 の情報処理装置が通信可能に接続され、
前記第 1 の情報処理装置は、
前記第 1 の記憶ボリュームへの第 2 のハートビートメッセージの書き込み要求を前記第 1 のストレージ装置に所定時間以内の間隔で繰り返し送信するノードハートビート書き込み要求部
を備え、
前記第 1 のストレージ装置は、
前記第 2 のハートビートメッセージの書き込み要求に応じて、前記第 2 のハートビートメッセージを前記第 1 の記憶ボリュームに書き込むノードハートビート書き込み部
を備え、
前記第 2 のストレージ装置は、
前記複製データ受信部により前記第 2 の記憶ボリュームへ書き込まれる前記第 2 のハートビートメッセージの複製を前記第 2 の情報処理装置に送信するノードハートビート送信部
を備え、
前記第 2 の情報処理装置は、
前記ノードハートビート送信部により送信される前記第 2 のハートビートメッセージの複製の検知を行うノードハートビート検知部
を備えることを特徴とする請求項 1 に記載のストレージシステム。

【請求項 3】

前記第 1 のストレージ装置は、
前記第 1 のハートビートメッセージであるディスクハートビート信号を作成するディスクハートビート作成部
を備えることを特徴とする請求項 1 に記載のストレージシステム。

【請求項 4】

前記第 1 の情報処理装置は、

前記第2のハートビートメッセージであるノードハートビート信号を作成するノードハートビート作成部
を備えることを特徴とする請求項2に記載のストレージシステム。

【請求項5】

前記第2のストレージ装置は、
前記ディスクハートビート検知部による前記第1のハートビートメッセージの複製の検知の結果を前記第2の情報処理装置に送信するディスクハートビート検知結果送信部を備え、
前記第2の情報処理装置は、
前記ディスクハートビート検知結果送信部により送信される前記第1のハートビートメッセージの複製の検知の結果と、前記ノードハートビート検知部による前記第2のハートビートメッセージの複製の検知の結果とに応じて、前記第1の情報処理装置及び前記第1のストレージ装置を含んで構成される第1のコンピュータシステムの動作状態を判定する動作状態判定部
を備えることを特徴とする請求項2に記載のストレージシステム。

【請求項6】

前記第2の情報処理装置は、
前記第1のコンピュータシステムの動作状態に応じて、前記第2の情報処理装置及び前記第2のストレージ装置を含んで構成される第2のコンピュータシステムに、前記第1のコンピュータシステムにおける情報処理をフェイルオーバーさせるフェイルオーバー制御部を備えることを特徴とする請求項5に記載のストレージシステム。

【請求項7】

前記第2の情報処理装置は、
前記第1のコンピュータシステムの動作状態をユーザインタフェースに出力する動作状態表示部
を備えることを特徴とする請求項5に記載のストレージシステム。

【請求項8】

前記第1のハートビートメッセージは、
前記第1のハートビートメッセージの識別情報、
前記第1のハートビートメッセージが作成された時刻を示す情報、
前記第1のハートビートメッセージが書き込まれる前記第1の記憶ボリュームの記憶位置を示す情報、
及び、前記第1のハートビートメッセージが書き込まれる前記第2の記憶ボリュームの記憶位置を示す情報、
の少なくともいずれかを含んで構成されること
を特徴とする請求項1に記載のストレージシステム。

【請求項9】

データを記憶する第1の記憶ボリュームを有する第1のストレージ装置と、
前記第1のストレージ装置と通信可能に接続され、データを記憶する第2の記憶ボリュームを有する第2のストレージ装置と、
を備え、
前記第1のストレージ装置は、
前記第1の記憶ボリュームにデータが書き込まれると、前記データの複製を前記第2のストレージ装置に送信する複製データ送信部
を備え、
前記第2のストレージ装置は、
前記複製データ送信部により送信される前記データの複製を受信し、前記データの複製を前記第2の記憶ボリュームに書き込む複製データ受信部
を備えるストレージシステムの制御方法であって、
前記第1のストレージ装置が、第1のハートビートメッセージを前記第1の記憶ボリ

ームに所定時間以内の間隔で繰り返し書き込み、

前記第2のストレージ装置が、前記複製データ受信部により前記第2の記憶ボリュームへ書き込まれる前記第1のハートビートメッセージの複製の検知を行うこと
を特徴とするストレージシステムの制御方法。

【請求項10】

前記第1のストレージ装置には第1の情報処理装置が通信可能に接続され、
前記第2のストレージ装置には第2の情報処理装置が通信可能に接続され、
前記第1の情報処理装置が、前記第1のストレージ装置に、前記第1の記憶ボリュームへの第2のハートビートメッセージの書き込み要求を所定時間以内の間隔で繰り返し送信し、
前記第1のストレージ装置が、前記第2のハートビートメッセージの書き込み要求に応じて、前記第2のハートビートメッセージを前記第1の記憶ボリュームに書き込み、
前記第2のストレージ装置が、前記複製データ受信部により前記第2の記憶ボリュームへ書き込まれる前記第2のハートビートメッセージの複製を前記第2の情報処理装置に送信し、

前記第2の情報処理装置が、前記第2のストレージ装置から送信される前記第2のハートビートメッセージの複製の検知を行うこと
を特徴とする請求項9に記載のストレージシステムの制御方法。

【請求項11】

前記第1のストレージ装置は、
前記第1のハートビートメッセージであるディスクハートビート信号を作成するディスクハートビート作成部
を備えることを特徴とする請求項9に記載のストレージシステムの制御方法。

【請求項12】

前記第1の情報処理装置は、
前記第2のハートビートメッセージであるノードハートビート信号を作成するノードハートビート作成部
を備えることを特徴とする請求項10に記載のストレージシステムの制御方法。

【請求項13】

前記第2のストレージ装置が、前記第1のハートビートメッセージの複製の検知の結果を前記第2の情報処理装置に送信し、
前記第2の情報処理装置が、前記第2のストレージ装置から送信される前記第1のハートビートメッセージの複製の検知の結果と、前記第2のハートビートメッセージの複製の検知の結果とに応じて、前記第1の情報処理装置及び前記第1のストレージ装置を含んで構成される第1のコンピュータシステムの動作状態を判定すること
を特徴とする請求項10に記載のストレージシステムの制御方法。

【請求項14】

前記第2の情報処理装置が、前記第1のコンピュータシステムの動作状態に応じて、前記第2の情報処理装置及び前記第2のストレージ装置を含んで構成される第2のコンピュータシステムに、前記第1のコンピュータシステムにおける情報処理をフェイルオーバーさせること
を特徴とする請求項13に記載のストレージシステムの制御方法。

【請求項15】

前記第2の情報処理装置が、前記第1のコンピュータシステムの動作状態をユーザインタフェースに出力すること
を特徴とする請求項13に記載のストレージシステムの制御方法。

【請求項16】

前記第1のハートビートメッセージは、
前記第1のハートビートメッセージの識別情報、
前記第1のハートビートメッセージが作成された時刻を示す情報、

前記第1のハートビートメッセージが書き込まれる前記第1の記憶ボリュームの記憶位置を示す情報、

及び、前記第1のハートビートメッセージが書き込まれる前記第2の記憶ボリュームの記憶位置を示す情報、

の少なくともいずれかを含んで構成されること

を特徴とする請求項9に記載のストレージシステムの制御方法。

【請求項17】

データを記憶する第1の記憶ボリュームを有する第1のストレージ装置、及び前記第1のストレージ装置と通信可能に接続される第1の情報処理装置を備える第1のコンピュータシステムと、

前記第1のストレージ装置と通信可能に接続され、データを記憶する第2の記憶ボリュームを有する第2のストレージ装置、及び前記第2のストレージ装置と通信可能に接続される第2の情報処理装置を備える第2のコンピュータシステムと、

を備え、

前記第1のストレージ装置は、

前記第1の記憶ボリュームにデータが書き込まれると、前記データの複製を前記第2のストレージ装置に送信する複製データ送信部

を備え、

前記第2のストレージ装置は、

前記複製データ送信部により送信される前記データの複製を受信し、前記データの複製を前記第2の記憶ボリュームに書き込む複製データ受信部

を備え、

前記第1のストレージ装置は、

所定時間以内の間隔で繰り返し第1のハートビートメッセージを作成するディスクハートビート作成部と、

前記第1のハートビートメッセージを前記第1の記憶ボリュームに所定時間以内の間隔で繰り返し書き込むディスクハートビート書き込み部と

を備え、

前記第2のストレージ装置は、

前記複製データ受信部により前記第2の記憶ボリュームへ書き込まれる前記第1のハートビートメッセージの複製の検知を行うディスクハートビート検知部と、

前記ディスクハートビート検知部による前記第1のハートビートメッセージの複製の検知の結果を前記第2の情報処理装置に送信するディスクハートビート検知結果送信部と、

を備え、

前記第1の情報処理装置は、

所定時間以内の間隔で繰り返し第2のハートビートメッセージを作成するノードハートビート作成部と、

前記第1の記憶ボリュームへの前記第2のハートビートメッセージの書き込み要求を前記第1のストレージ装置に所定時間以内の間隔で繰り返し送信するノードハートビート書き込み要求部と

を備え、

前記第1のストレージ装置は、

前記第2のハートビートメッセージの書き込み要求に応じて、前記第2のハートビートメッセージを前記第1の記憶ボリュームに書き込むノードハートビート書き込み部

を備え、

前記第2のストレージ装置は、

前記複製データ受信部により前記第2の記憶ボリュームへ書き込まれる前記第2のハートビートメッセージの複製を前記第2の情報処理装置に送信するノードハートビート送信部

を備え、

前記第 2 の情報処理装置は、

前記ノードハートビート送信部により送信される前記第 2 のハートビートメッセージの複製の検知を行うノードハートビート検知部と、

前記ノードハートビート検知部による前記第 2 のハートビートメッセージの複製の検知の結果と、前記ディスクハートビート検知結果送信部により送信される前記第 1 のハートビートメッセージの複製の検知の結果とに応じて、前記第 1 のコンピュータシステムの動作状態を判定する動作状態判定部と、

前記第 1 のコンピュータシステムの動作状態に応じて、前記第 2 のコンピュータシステムに、前記第 1 のコンピュータシステムにおける情報処理をフェイルオーバーさせるフェイルオーバー実行部と、

を備えること

を特徴とするストレージシステム。

【請求項 18】

前記第 1 のハートビートメッセージは、

前記第 1 のハートビートメッセージの識別情報、

前記第 1 のハートビートメッセージが作成された時刻を示す情報、

前記第 1 のハートビートメッセージが書き込まれる前記第 1 の記憶ボリュームの記憶位置を示す情報、

及び、前記第 1 のハートビートメッセージが書き込まれる前記第 2 の記憶ボリュームの記憶位置を示す情報、

の少なくともいずれかを含んで構成され、

前記第 2 のハートビートメッセージは、

前記第 2 のハートビートメッセージの識別情報、

前記第 2 のハートビートメッセージが作成された時刻を示す情報、

前記第 2 のハートビートメッセージが書き込まれる前記第 1 の記憶ボリュームの記憶位置を示す情報、

及び、前記第 2 のハートビートメッセージが書き込まれる前記第 2 の記憶ボリュームの記憶位置を示す情報、

の少なくともいずれかを含んで構成されること

を特徴とする請求項 17 に記載のストレージシステム。

【請求項 19】

データを記憶する第 2 の記憶ボリュームに対するデータの読み書きを制御する第 2 のストレージ制御装置と通信可能に接続され、データを記憶する第 1、第 3 の記憶ボリュームに対するデータの読み書きを制御する第 1 のストレージ制御装置であって、

前記第 1 の記憶ボリュームにデータが書き込まれると、当該データの複製を前記第 2 のストレージ制御装置に送信する複製データ送信部と、

第 1 のハートビートメッセージを前記第 1 の記憶ボリュームに書き込むディスクハートビート書き込み部と、

前記第 2 のストレージ制御装置により前記第 2 の記憶ボリュームに書き込まれた第 2 のハートビートメッセージの複製を、前記第 2 のストレージ制御装置から受信し、当該複製を前記第 3 の記憶ボリュームに書き込む複製データ受信部と、

前記複製データ受信部により前記第 3 の記憶ボリュームへ書き込まれる前記第 2 のハートビートメッセージの複製の検知を行うディスクハートビート検知部と、
を備えることを特徴とするストレージ制御装置。

【書類名】 明細書

【発明の名称】 ストレージシステム、ストレージシステムの制御方法、及びストレージ制御装置

【技術分野】**【0001】**

本発明は、ストレージシステム、ストレージシステムの制御方法、及びストレージ制御装置に関する。

【背景技術】**【0002】**

近年の情報技術の進歩により、大地震等の災害発生時にコンピュータシステムを構成するストレージ装置に記憶されるデータの消失を回避するため、レプリケーションと呼ばれる技術が用いられるようになっている。レプリケーションは、メインサイトのコンピュータシステムが備えるストレージ装置の記憶ボリュームに記憶されるデータの複製を、リモートサイトのコンピュータシステムが備えるストレージ装置の記憶ボリュームに記憶するための技術である。

【0003】

また、メインサイトのコンピュータシステムが停止したことをリモートサイトのコンピュータシステムが検知した場合には、リモートサイトのコンピュータシステムが、それまでメインサイトのコンピュータシステムにより行われていた情報処理を引き継いで行う技術も開発されている。

【0004】

メインサイトのコンピュータシステムが備える情報処理装置は、ハートビートメッセージと呼ばれる所定のデータをリモートサイトのコンピュータシステムに所定時間以内の間隔で繰り返し送信する。このハートビートメッセージの送信を、レプリケーションの技術を利用して行う技術も開発されている。

【特許文献1】 特開 2002-312189 号公報

【発明の開示】**【発明が解決しようとする課題】****【0005】**

この場合、リモートサイトのコンピュータシステムは、メインサイトからのハートビートメッセージの送信が途絶えたことを検知することにより、メインサイトのコンピュータシステムに異常が発生したことを検知することができる。しかしながら、メインサイトのストレージ装置の異常によりメインサイトからのハートビートメッセージの送信が途絶えた場合に、リモートサイトのコンピュータシステムは、異常が発生しているのがメインサイトのストレージ装置であることを特定することができない。

【0006】

本発明は上記課題を鑑みてなされたものであり、ストレージシステム、ストレージシステムの制御方法、及びストレージ制御装置を提供することを主たる目的とする。

【課題を解決するための手段】**【0007】**

上記課題を解決するために、本発明は、データを記憶する第1の記憶ボリュームを有する第1のストレージ装置と、前記第1のストレージ装置と通信可能に接続され、データを記憶する第2の記憶ボリュームを有する第2のストレージ装置と、を備え、前記第1のストレージ装置は、前記第1の記憶ボリュームにデータが書き込まれると、前記データの複製を前記第2のストレージ装置に送信する複製データ送信部を備え、前記第2のストレージ装置は、前記複製データ送信部により送信される前記データの複製を受信し、前記データの複製を前記第2の記憶ボリュームに書き込む複製データ受信部を備え、前記第1のストレージ装置は、第1のハートビートメッセージを前記第1の記憶ボリュームに所定時間以内の間隔で繰り返し書き込むディスクハートビート書き込み部を備え、前記第2のストレージ装置は、前記複製データ受信部により前記第2の記憶ボリュームへ書き込まれる前

記第1のハートビートメッセージの複製の検知を行うディスクハートビート検知部を備えることを特徴とするストレージシステムに関する。

【0008】

ここで、前記第1のストレージ装置には第1の情報処理装置が通信可能に接続され、前記第2のストレージ装置には第2の情報処理装置が通信可能に接続され、前記第1の情報処理装置は、前記第1の記憶ボリュームへの第2のハートビートメッセージの書き込み要求を前記第1のストレージ装置に所定時間以内の間隔で繰り返し送信するノードハートビート書き込み要求部を備え、前記第1のストレージ装置は、前記第2のハートビートメッセージの書き込み要求に応じて、前記第2のハートビートメッセージを前記第1の記憶ボリュームに書き込むノードハートビート書き込み部を備え、前記第2のストレージ装置は、前記複製データ受信部により前記第2の記憶ボリュームへ書き込まれる前記第2のハートビートメッセージの複製を前記第2の情報処理装置に送信するノードハートビート送信部を備え、前記第2の情報処理装置は、前記ノードハートビート送信部により送信される前記第2のハートビートメッセージの複製の検知を行うノードハートビート検知部を備えるようにすることもできる。

【0009】

なお、以下第1のハートビートメッセージをディスクハートビート信号、第2のハートビートメッセージをノードハートビート信号とも記す。

【0010】

その他、本願が開示する課題、及びその解決方法は、発明を実施するための最良の形態の欄、及び図面により明らかにされる。

【発明の効果】

【0011】

ストレージシステム、ストレージシステムの制御方法、及びストレージ制御装置を提供することができる。

【発明を実施するための最良の形態】

【0012】

===全体構成例===

まず、本実施の形態に係るストレージシステム400の全体構成を示すブロック図を図1に示す。

本実施の形態に係るストレージシステム400は、情報処理装置A（第1の情報処理装置）（100）及びストレージ装置A（第1のストレージ装置）（200）と、情報処理装置B（第2の情報処理装置）（100）及びストレージ装置B（第2のストレージ装置）（200）とを含んで構成される。

【0013】

情報処理装置A（100）とストレージ装置A（200）とは、メインサイト（プライマリサイト）に設置されるコンピュータシステム（第1のコンピュータシステム）を構成する。また情報処理装置B（100）とストレージ装置B（200）とは、リモートサイト（セカンダリサイト）に設置されるコンピュータシステム（第2のコンピュータシステム）を構成する。

【0014】

情報処理装置A（100）は、ストレージ装置A（200）が備える記憶ボリューム230により提供される記憶資源を利用し、各種の情報処理サービスを提供するための情報機器である。情報処理装置A（100）により提供される情報処理サービスとしては、例えば銀行の自動預金預け払いサービスや、航空機の座席予約サービス等とすることができる。情報処理装置A（100）とストレージ装置A（200）とが設置される場所をメインサイトと呼ぶ。

【0015】

記憶ボリューム230は、ハードディスクドライブ等により提供される物理的な記憶領域である物理ボリュームと、物理ボリューム上に論理的に設定される記憶領域である論理

ボリュームとを含む、データを記憶するための記憶領域である。

【0016】

一方、情報処理装置 B (100) は、災害の発生などによりメインサイトに設置される情報処理装置 A (100) による情報処理サービスの提供ができない状態になったときに、情報処理装置 A (100) が行っていた情報処理サービスの提供を引き継いで行うための情報機器である。その際、情報処理装置 B (100) はストレージ装置 B (200) が備える記憶ボリューム 230 により提供される記憶資源を利用することにより情報処理サービスの提供を行う。そのためストレージ装置 A (200) が備える記憶ボリューム 230 に記憶されているデータの複製は、ストレージ装置 B (200) が備える記憶ボリューム 230 にも記憶されている。ストレージ装置 A (200) に記憶されているデータの複製をストレージ装置 B (200) に記憶するために、レプリケーションの制御が行われる。レプリケーションの制御が行われる場合には、ストレージ装置 A (200) の記憶ボリューム 230 にデータが書き込まれると、ストレージ装置 A (200) は、そのデータの複製をストレージ装置 B (200) に送信する。そしてストレージ装置 B (200) は、ストレージ装置 A (200) から送信されるそのデータの複製を受信し、そのデータの複製をストレージ装置 B (200) の記憶ボリューム 230 に書き込む。レプリケーションの制御については後述する。情報処理装置 B (100) とストレージ装置 B (200) とが設置される場所をリモートサイトと呼ぶ。

【0017】

なお情報処理装置 B (100) は、情報処理装置 A (100) により提供される情報処理サービスを引き継いで行うだけでなく、独自の情報処理サービスの提供を行うようにすることもできる。この場合は、災害の発生により、情報処理装置 B (100) による情報処理サービスの提供が行えない状態になった際に、情報処理装置 A (100) に情報処理サービスの提供を引き継いで行わせるようにすることもできる。この場合、2つのコンピュータシステムは、互いに自己のコンピュータシステムがメインサイトに設置されるコンピュータシステムとなり、相手側のコンピュータシステムがリモートサイトに設置されるコンピュータシステムとなる。ここでは説明の簡略化のため、上述したように情報処理装置 A (100) とストレージ装置 A (200) とが設置される場所をメインサイトとし、情報処理装置 B (100) とストレージ装置 B (200) とが設置される場所をリモートサイトとする。

【0018】

メインサイトに設置される情報処理装置 A (100) と、リモートサイトに設置される情報処理装置 B (100) とは、第1のネットワーク 300 で通信可能に接続されている。また、メインサイトに設置されるストレージ装置 A (200) と、リモートサイトに設置されるストレージ装置 B (200) とは、第2のネットワーク 310 で通信可能に接続されている。上述したレプリケーションの制御は第2のネットワーク 310 を介してデータの複製が送信されることにより行われる。また情報処理装置 A (100) とストレージ装置 A (200) との間は、入出力バス A (330) で通信可能に接続されている。情報処理装置 B (100) とストレージ装置 B (200) との間も、入出力バス B (330) で通信可能に接続されている。

【0019】

第1のネットワーク 300 としては、例えば TCP/IP (Transmission Control Protocol / Internet Protocol) 通信プロトコルにより通信が行われる LAN (Local Area Network) や WAN (Wide Area Network) とすることができる。

【0020】

また第2のネットワーク 310 としては、例えばファイバチャネル通信プロトコルにより通信が行われる SAN (Storage Area Network) とすることができる。入出力バス 330 も、例えばファイバチャネル通信プロトコルにより通信が行われる SAN とすることができる。もちろん、例えば FICON (Fibre Connection) (登録商標) や ESCON (Enterprise System Connection) (登録商標)、ACONARC (Advanced Connection A

rchitecture) (登録商標)、F I B A R C (Fibre Connection Architecture) (登録商標)、i S C S I (Internet Small Computer Systems Interface) 等の通信プロトコルを用いるようにすることもできる。ストレージ装置 A (200) とストレージ装置 B (200) との間を、このような信頼性の高い通信プロトコルを用いて第2のネットワーク 310 で接続することにより、メインサイトのストレージ装置 A (200) とリモートサイトのストレージ装置 B (200) との間のデータの送受信を高信頼に行うことが可能となる。

【0021】

=== 情報処理装置 ===

次に、本実施の形態に係る情報処理装置 100 の構成を示すブロック図を図2に示す。なお本実施の形態に係る情報処理装置 100 には、情報処理装置 A (100) と情報処理装置 B (100) とがあるが、いずれも基本的な構成は同様であるため、特に両者を区別する必要がない場合には、情報処理装置 100 と記す。

【0022】

情報処理装置 100 は、CPU 110、メモリ 120、ポート 130、記録媒体読取装置 140、入力装置 150、出力装置 160、記憶装置 170 を備える。

【0023】

CPU 110 は情報処理装置 100 の全体の制御を司るもので、メモリ 120 に記憶されたアプリケーションプログラム 190 を実行することにより上述した各種情報処理サービスの提供を行う。また CPU 110 は、メモリ 120 に記憶された、本実施の形態に係る各種の動作を行うためのコードから構成されるハートビート制御プログラム 191 や、ノードハートビート制御プログラム 192、ストレージ管理プログラム 193、クラスタリング制御プログラム 196、オペレーティングシステム 197 を実行する。

例えば CPU 110 によりハートビート制御プログラム 191 が実行されることにより、動作状態判定部や動作状態表示部が実現される。また例えば CPU 110 によりノードハートビート制御プログラム 192 が実行されることにより、ノードハートビート書き込み要求部や、ノードハートビート検知部、ノードハートビート作成部が実現される。また例えば CPU 110 によりクラスタリング制御プログラム 196 が実行されることにより、フェイルオーバー制御部が実現される。またメモリ 120 には、ハートビート状態テーブル 194、活性化／不活性化メッセージ 195、ハートビート判定テーブル 198 が記憶される。これらの詳細は後述する。なお、ハートビート制御プログラム 191 や、ノードハートビート制御プログラム 192、ストレージ管理プログラム 193、クラスタリング制御プログラム 196、オペレーティングシステム 197 は、それぞれ個別のプログラムとすることもできるし、これらのプログラムの少なくとも一部が同一プログラムにより構成されるようにすることもできる。また各プログラムが複数のプログラムにより構成されるようにすることもできる。

【0024】

記録媒体読取装置 140 は、記録媒体 170 に記録されているプログラムやデータを読み取るための装置である。読み取られたプログラムやデータはメモリ 120 や記憶装置 170 に格納される。従って、例えば記録媒体 170 に記録されたハートビート制御プログラム 191 や、ノードハートビート制御プログラム 192、ストレージ管理プログラム 193、クラスタリング制御プログラム 196、オペレーティングシステム 197 を、記録媒体読取装置 140 を用いて上記記録媒体 170 から読み取って、メモリ 120 や記憶装置 180 に記憶するようにすることができる。記録媒体 170 としてはフレキシブルディスクや磁気テープ、CD-ROM、半導体メモリ等を用いることができる。記録媒体読取装置 140 は情報処理装置 100 に内蔵されている形態とすることもできるし、外付されている形態とすることもできる。記憶装置 180 は、例えばハードディスク装置や半導体記憶装置等である。ハートビート制御プログラム 191 や、ノードハートビート制御プログラム 192、ストレージ管理プログラム 193、クラスタリング制御プログラム 196、オペレーティングシステム 197、ハートビート状態テーブル 194、活性化／不活性

化メッセージ195、ハートビート判定テーブル198は、記憶装置180に記憶されるようにすることもできる。

入力装置150はオペレータ等による情報処理装置100へのデータ入力等のために用いられるユーザインタフェースである。入力装置150としては例えばキーボードやマウス等を用いることができる。出力装置160は情報を外部に出力するためのユーザインタフェースである。出力装置160としては例えばディスプレイやプリンタ等を用いることができる。ポート130は通信を行うための装置である。例えば第1のネットワーク300を介して行う他の情報処理装置100との通信や、ストレージ装置200に対するデータ入出力要求の送信などは、ポート130を介して行われるようにすることができる。また例えば、アプリケーションプログラム190や、ハートビート制御プログラム191、ノードハートビート制御プログラム192、ストレージ管理プログラム193、クラスタリング制御プログラム196、オペレーティングシステム197を、ポート130を通じて他の情報処理装置100から受信して、メモリ120や記憶装置180に記憶するようにすることもできる。

【0025】

===ストレージ装置===

次に本実施の形態に係るストレージ装置200について、図1を用いて説明する。なお本実施の形態に係るストレージ装置200には、ストレージ装置A(200)とストレージ装置B(200)とがあるが、いずれも基本的な構成は同様であるため、特に両者を区別する必要がない場合には、ストレージ装置200と記す。

【0026】

ストレージ装置200は、ディスク制御装置(ストレージ制御装置)210と、ディスク駆動装置220と、管理端末260とを備える。

ディスク制御装置210は、情報処理装置100や他のストレージ装置200との間で通信を行い、ディスク駆動装置220が備える記憶ボリューム230に対するデータの読み書きを制御する。例えば情報処理装置100からデータの書き込み要求を受信し、ディスク駆動装置220が備える記憶ボリューム230に対するデータの書き込み処理を行う。

ディスク制御装置210は、CPU211、メモリ212、FC I/F(Fibre Channel InterFace)213、ディスクI/F(ディスクInterFace)214を備える。

【0027】

CPU211は、ディスク制御装置210の全体の制御を司るもので、メモリ212に記憶された、本実施の形態に係る各種の動作を行うためのコードから構成されるディスクハートビート制御プログラム215や、レプリケーション制御プログラム217、ストレージ制御プログラム218を実行する。

例えばCPU211によりディスクハートビート制御プログラム215が実行されることにより、ディスクハートビート書き込み部やディスクハートビート検知部、ディスクハートビート作成部が実現される。また例えばCPU211によりレプリケーション制御プログラム217が実行されることにより、複製データ送信部や複製データ受信部が実現される。また例えばCPU211によりストレージ制御プログラム218が実行されることにより、ノードハートビート書き込み部や、ノードハートビート送信部、ディスクハートビート検知結果送信部が実現される。またメモリ212にはペア管理テーブル216が記憶されている。詳細は後述する。なお、ディスクハートビート制御プログラム215や、レプリケーション制御プログラム217、ストレージ制御プログラム218は、それぞれ個別のプログラムとすることもできるし、これらのプログラムの少なくとも一部が同一プログラムにより構成されるようにすることもできる。また各プログラムが複数のプログラムにより構成されるようにすることもできる。FC I/F213は情報処理装置100や他のストレージ装置200と通信を行うためのインタフェース機能を有する。ディスクI/F214はディスク駆動装置220とデータの授受を行うためのインタフェース機能を有する。

【0028】

なお、第2のネットワーク310で通信可能に接続されたディスク制御装置A(210)とディスク制御装置B(210)とを、ストレージ制御システムとも記す。

【0029】

ディスク駆動装置220はデータを記憶する記憶ボリューム230を有する。記憶ボリューム230は、ハードディスクドライブ等により提供される物理的な記憶領域である物理ボリュームと、物理ボリューム上に論理的に設定される記憶領域である論理ボリュームとを含む記憶領域である。図1では、ストレージ装置A(200)においてはLU01、LU02、LU03の3つの記憶ボリューム(第1の記憶ボリューム、第3の記憶ボリューム)230が示され、ストレージ装置B(200)においてはLU11、LU12、LU13の3つの記憶ボリューム(第2の記憶ボリューム)230が示されている。

【0030】

管理端末260は、ストレージ装置200の保守管理を行うための情報機器である。管理端末260はストレージ装置200に内蔵されるように構成することもできるし、外付けされるように構成することもできる。また管理端末260はストレージ装置200の保守・管理を専用に行うコンピュータとすることもできるし、汎用のコンピュータに保守・管理機能を持たせたものとすることもできる。

【0031】

管理端末260の構成を示すブロック図を図3に示す。

管理端末260は、CPU261、メモリ262、ポート263、記録媒体読取装置264、入力装置265、出力装置266、記憶装置268を備える。

【0032】

CPU261は管理端末260の全体の制御を司るもので、メモリ262に格納された、本実施の形態に係る各種の動作を行うためのコードから構成される制御プログラム269を実行することによりストレージ装置200の保守管理機能等を実現する。

記録媒体読取装置264は、記録媒体267に記録されているプログラムやデータを読み取るための装置である。読み取られたプログラムやデータはメモリ262や記憶装置268に格納される。従って、例えば記録媒体267に記録された制御プログラム269やディスクハートビート制御プログラム215、レプリケーション制御プログラム217、ストレージ制御プログラム218を、記録媒体読取装置264を用いて上記記録媒体267から読み取って、メモリ262や記憶装置268に記憶することができる。記録媒体267としてはフレキシブルディスクやCD-ROM、磁気テープ、半導体メモリ等を用いることができる。記録媒体読取装置264は管理端末260に内蔵されている形態とすることもできるし、外付けされている形態とすることもできる。記憶装置268は、例えばハードディスク装置や半導体記憶装置等とすることができる。制御プログラム269は記憶装置268に記憶されるようにすることもできる。

入力装置265はオペレータ等による管理端末260へのデータ入力等のために用いられるユーザインタフェースである。入力装置265としては例えばキーボードやマウス等を用いることができる。出力装置266は情報を外部に出力するためのユーザインタフェースである。出力装置266としては例えばディスプレイやプリンタ等を用いることができる。ポート263は情報処理装置100やストレージ装置200と通信を行うための装置である。これにより例えば、制御プログラム269やディスクハートビート制御プログラム215、レプリケーション制御プログラム217、ストレージ制御プログラム218を、情報処理装置100からポート263を通じて受信してメモリ262や記憶装置268に記憶することができる。またメモリ262や記憶装置268に記憶されたディスクハートビート制御プログラム215やレプリケーション制御プログラム217、ストレージ制御プログラム218を、ポート263を通じてディスク制御装置210に送信して、ディスク制御装置210のメモリ212に記憶させるようにすることもできる。

【0033】

なお、本実施の形態に係る管理端末260の機能を情報処理装置100に設けるように

することもできる。この場合、管理端末 260 を設けない構成とすることもできる。またこの場合、ストレージ装置 200 の保守管理は情報処理装置 100 を用いて行われることになる。

【0034】

===レプリケーション制御===

次に本実施の形態に係るストレージシステム 400 において行われるレプリケーションの制御について説明する。レプリケーションの制御は、ディスク制御装置 210 の CPU 211 が、レプリケーション制御プログラム 217 を実行することにより行うことができる。レプリケーションの対象となる記憶ボリューム 230 や、レプリケーションの制御の方式等は、ペア管理テーブル 216 に定義されている。ペア管理テーブル 216 は、例えば情報処理装置 100 においてストレージ管理プログラム 193 が実行されることにより、情報処理装置 100 を操作するオペレータ等により入力装置 150 から入力されたデータに基づいて、作成されるようにすることができる。なお、ストレージ管理プログラム 193 を管理端末 260 に備える様にし、管理端末 260 の CPU 261 にストレージ管理プログラム 193 を実行させるようにすることにより、管理端末 260 を操作するオペレータ等により入力装置 265 から入力されたデータに基づいて、ペア管理テーブル 216 が作成されるようにすることもできる。

【0035】

ペア管理テーブル 216 を図 4 に示す。

ペア管理テーブル 216 は、「ペア種類」欄、「複製方式」欄、「複製元装置」欄、「複製先装置」欄、「複製元ボリューム」欄、「複製先ボリューム」欄、「ペア状態」欄を備える。

【0036】

ペアとは 2 つの記憶ボリューム 230 により形成される記憶ボリューム 230 の組み合わせを言う。また、ペアを形成する 2 つの記憶ボリューム 230 が同一のストレージ装置 200 にある場合を「ローカルペア」、異なるストレージ装置 200 にある場合を「リモートペア」と記す。ペアを形成する記憶ボリューム 230 は、一方を主記憶ボリューム 230、他方を副記憶ボリューム 230 として管理される。主記憶ボリューム 230 にデータが書き込まれるとそのデータの複製が副記憶ボリューム 230 に書き込まれる。一つの主記憶ボリューム 230 に対して複数の副記憶ボリューム 230 を組み合わせることも可能である。

【0037】

ストレージ管理プログラム 193 を実行する情報処理装置 200 が、データの複製元となるストレージ装置 A (200) にリモートペアの形成を指示すると、ストレージ装置 A (200) はペア管理テーブル 216 を更新する。その後ストレージ装置 A (200) は、データの複製先となるストレージ装置 B (200) にリモートペアの形成を指示する。そしてストレージ装置 B (200) は、ストレージ装置 B (200) 内のメモリ 212 のペア管理テーブル 216 を更新する。

【0038】

ペア管理テーブル 216 の「ペア種類」欄は当該ペアがローカルペアであるかリモートペアであるかを示す。「複製方式」欄は、当該ペアがリモートペアである場合に、レプリケーションの方式が同期方式であるか非同期方式であるかを示す。「複製元装置」欄と「複製先装置」欄は、当該ペアがリモートペアである場合に、複製元のストレージ装置 200 と複製先のストレージ装置 200 を示す。「複製元ボリューム」欄は当該ペアの主記憶ボリューム 230 の識別番号を、「複製先ボリューム」欄は当該ペアの副記憶ボリューム 230 の識別番号を示す。記憶ボリューム 230 の識別番号としては、例えば LUN (Logical Unit Number) とすることができる。

【0039】

「ペア状態」欄は当該ペアの状態を示す。ペアの状態としては、「ペア中」、「スプリット中」、及び「リシンク中」がある。

「ペア中」の場合は、情報処理装置 100 から主記憶ボリューム 230 に書き込まれたデータの複製は副記憶ボリューム 230 にも書き込まれる。このような主記憶ボリューム 230 と副記憶ボリューム 230 との対応付けにより、主記憶ボリューム 230 に記憶されている内容と副記憶ボリューム 230 に記憶されている内容との同一性を確保することができる。

「スプリット中」の場合は、情報処理装置 100 から主記憶ボリューム 230 にデータが書き込まれても副記憶ボリューム 230 には反映されない。

「リシンク中」は「スプリット中」から「ペア中」に移行する途中の状態である。すなわち、「スプリット中」に主記憶ボリューム 230 に対してなされたデータの更新を副記憶ボリューム 230 に反映している状態である。反映が完了すると当該ペアの状態は「ペア中」になる。

【0040】

上記ペアの形成、ペアのスプリット、及びペアのリシンクは、ストレージ制御プログラム 193 が実行されている情報処理装置 100 に対して入力装置 150 からオペレータが指示入力を与えることにより行われるようにすることができる。オペレータにより与えられた指示入力はディスク制御装置 210 に送信される。ディスク制御装置 210 は、レプリケーション制御プログラム 217 を実行して、上記指示に従って、ペアの形成やペア状態の変更を行う。ディスク制御装置 210 は、形成したペアのペア状態に応じて、例えば「ペア中」の主記憶ボリューム 230 に対してデータの書き込み要求を情報処理装置 100 から受信した場合には、データを主記憶ボリューム 230 に書き込むと共に、ペア管理テーブル 216 を参照して、そのデータの複製が書き込まれる記憶ボリューム B (230) を特定し、そのデータの複製をストレージ装置 B (200) に送信する。そしてストレージ装置 B (200) は、送信されたこのデータの複製を受信し、そのデータの複製を記憶ボリューム B (230) に書き込む。

【0041】

以上のレプリケーションの制御を行うことによって、メインサイトのコンピュータシステムのデータを、リモートサイトのコンピュータシステムに記憶しておくようにすることが可能となる。

【0042】

=== クラスタ制御 ===

次に、本実施の形態に係るストレージシステム 400 で行われるクラスタ制御について説明する。

【0043】

本実施の形態に係る情報処理装置 A (100) 及び情報処理装置 B (100) は、互いにクラスタリング制御プログラム 196 を実行することにより、クラスタ制御を行う。クラスタ制御とは、メインサイトのコンピュータシステムが停止したことをリモートサイトのコンピュータシステムが検知した場合には、リモートサイトのコンピュータシステムが、それまでメインサイトのコンピュータシステムにより行われていた情報処理を引き継いで行うための制御をいう。なお情報処理を引き継ぐことをフェイルオーバーともいう。

【0044】

ここで、メインサイトのコンピュータシステムが停止したことをリモートサイトのコンピュータシステムが検知するために、通常、メインサイトのコンピュータシステムは、リモートサイトのコンピュータシステムに対して、ハートビート信号（ハートビートメッセージ）と呼ばれる所定のデータを、所定時間以内の間隔で繰り返し送信する。リモートサイトのコンピュータシステムは、メインサイトのコンピュータシステムから送信されるハートビート信号を検知することにより、メインサイトのコンピュータシステムの動作状態を判定することができる。例えばリモートサイトのコンピュータシステムは、所定時間以上待ってもメインサイトのコンピュータシステムからのハートビート信号を検知できない場合には、メインサイトのコンピュータシステムに異常が発生したと判定することができる。

【0045】

===ハートビート信号の送受信===

次に本実施の形態に係るストレージシステム400において行われるハートビート信号の送受信の制御について、図5乃至図14を用いて説明する。

【0046】

まず本実施の形態に係るハートビート信号の送受信の制御を説明するためのシステム構成図を図5に示す。図5に示すシステム構成図は、図1に示したシステム構成図と比較して、メインサイト及びリモートサイトのそれぞれにおいて、複数の情報処理装置100を備えている点異なる。この場合、例えばメインサイトのそれぞれの情報処理装置A(100)間において互いにハートビート信号の授受を行うことにより、メインサイト内の情報処理装置A(100)においてクラスタ制御を行うようにすることができる。すなわち、メインサイト内のある情報処理装置A(100)が動作を停止したことを、メインサイト内の他の情報処理装置A(100)が検知した場合、それまでその情報処理装置A(100)が行っていた情報処理を、メインサイト内の他の情報処理装置A(100)が引き継いで行うようにすることができる。リモートサイトの情報処理装置B(100)についても同様である。

【0047】

また図5に示す「マイクロ制御VOL(VOLume)」、「ハートビートPVOL(Primary VOLume)」、「ユーザPVOL」、「ハートビートSVOL(Secondary VOLume)」、「ユーザSVOL」は、それぞれ記憶ボリューム230である。

マイクロ制御VOL(230)は、本実施の形態に係るハートビート信号の送受信の制御を行う際に用いられる記憶ボリューム230である。マイクロ制御VOL(230)には、ハートビート状態テーブル194が記憶される。詳細は後述する。

ハートビートPVOL(230)及びハートビートSVOL(230)は、レプリケーションのペアを構成しており、それぞれ主記憶ボリューム230及び副記憶ボリューム230である。ハートビートPVOL(230)には、ハートビート信号232が書き込まれる。そうすると、そのハートビート信号232の複製がペアを構成するハートビートSVOL(230)に書き込まれる。

ユーザPVOL(230)及びユーザSVOL(230)は、レプリケーションのペアを構成しており、それぞれ主記憶ボリューム230及び副記憶ボリューム230である。ユーザPVOL(230)には、情報処理装置100でアプリケーションプログラム190が実行されることにより生成される各種データが書き込まれる。そうすると、そのデータの複製が、ペアを構成するユーザSVOL(230)に書き込まれる。

【0048】

本実施の形態に係るハートビート信号232の送受信は、オペレーティングシステム197の制御の下、情報処理装置100でハートビート制御プログラム191及びノードハートビート制御プログラム192が実行され、ストレージ装置200でディスクハートビート制御プログラム215、ストレージ制御プログラム218、及びレプリケーション制御プログラム217が実行されることにより行われる。

【0049】

オペレーティングシステム197は、情報処理装置100でハートビート制御プログラム191等の各種プログラムが実行される際のAPI(Application Program Interface)を提供する。例えばオペレーティングシステム197は、記憶ボリューム230に対する"OPEN"、"READ"、"WRITE"、"CLOSE"を提供する。この場合、例えばノードハートビート制御プログラム192がストレージ装置(200)のハートビートPVOL(230)にハートビート信号232を書き込む場合には、ハートビートPVOL(230)へのポインタを得る際に"OPEN"を、ハートビート信号232を書き込む際に"WRITE"を、ハートビートPVOL(230)へのポインタを放棄する際に"CLOSE"を利用する。またノードハートビート制御プログラム192がストレージ装置(200)のハートビートSVOL(230)からハートビート信号232を

読み出す場合には、ハートビートSVOL (230) へのポインタを得る際に”OPEN”を、ハートビート信号232を読み出す際に”READ”を、ハートビートSVOL (230) へのポインタを放棄する際に”CLOSE”を利用する。

【0050】

ハートビート制御プログラム191は、情報処理装置100のメモリ120に記憶される活性化／不活性化メッセージ195やハートビート状態テーブル194の作成、更新、ストレージ装置200のマイクロ制御ボリューム230に記憶されるハートビート状態テーブル194の作成、更新等を行う。活性化／不活性化メッセージ195の作成や更新は、例えば情報処理装置100の入力装置150からオペレータによって入力されたデータに基づいて行われるようにすることができる。ハートビート状態テーブル194の作成、更新は、例えば活性化／不活性化メッセージ195の内容に基づいて行われるようにすることができる。またリモートサイトの情報処理装置100で実行されるハートビート制御プログラム191は、メインサイトのコンピュータシステムから送信されてくるハートビート信号232の検知の結果に応じて、メインサイトのコンピュータシステムの動作状態を判定する。詳細は後述する。

【0051】

ノードハートビート制御プログラム192は、メモリ120に記憶されたハートビート状態テーブル194に基づいて、ストレージ装置200に対して、所定時間以内の間隔で繰り返し、ハートビートPVOL (230) に対するハートビート信号 (第2のハートビートメッセージ、以下ノードハートビート信号232とも記す) 232の書き込み要求を送信する。このノードハートビート信号232は、ストレージ装置200によってハートビートPVOL (230) に書き込まれる。またノードハートビート制御プログラム192は、メモリ120に記憶されたハートビート状態テーブル194に基づいて、ストレージ装置200に対して、ハートビートSVOL (230) に書き込まれたノードハートビート信号232の複製の読み出し要求を送信する。このノードハートビート信号232の複製は、ストレージ装置200によってハートビートSVOL (230) から読みだされ、情報処理装置100に送信される。そしてノードハートビート制御プログラム192は、ストレージ装置200から送信されたノードハートビート信号232の複製の検知を行う。そしてメモリ120のハートビート状態テーブル194の内容を更新する。詳細は後述する。

【0052】

ディスクハートビート制御プログラム215は、マイクロ制御VOL (230) に記憶されたハートビート状態テーブル194に基づいて、所定時間以内の間隔で繰り返し、ハートビートPVOL (230) に対してハートビート信号 (第1のハートビートメッセージ、以下ディスクハートビート信号232とも記す) 232の書き込みを行う。ここでディスクハートビート制御プログラム215によってディスクハートビート信号232が書き込まれるハートビートPVOL (230) は、ノードハートビート制御プログラム192によって送信されるノードハートビート信号232が書き込まれるハートビートPVOL (230) と同じ記憶ボリューム230とすることもできるし、異なる記憶ボリューム230とすることもできる。またディスクハートビート制御プログラム215は、マイクロ制御VOL (230) に記憶されたハートビート状態テーブル194に基づいて、ハートビートSVOL (230) に書き込まれたディスクハートビート信号232の複製を読みだして、その検知を行う。そしてマイクロ制御VOL (230) に記憶されたハートビート状態テーブル194の内容を更新する。詳細は後述する。

【0053】

レプリケーション制御プログラム217は、ペア管理テーブル216に基づいて、ハートビートPVOL (230) にハートビート信号232が書き込まれると、そのハートビート信号232の複製を第2のネットワーク310を介してリモートサイトのストレージ装置200に送信する。またレプリケーション制御プログラム217は、第2のネットワーク310を介してメインサイトのストレージ装置200から送信されたハートビート信

号 232 の複製を受信し、そのハートビート信号 232 の複製をハートビート SVOL (230) に書き込む。

【0054】

ストレージ制御プログラム 218 は、情報処理装置 100 から送信されるノードハートビート信号 232 の書き込み要求に応じて、そのノードハートビート信号 232 をハートビート PVOL (230) に書き込む。また情報処理装置 100 から送信されるノードハートビート信号 232 の複製の読み出し要求に応じて、そのノードハートビート信号 232 の複製をハートビート SVOL (230) から読みだして、情報処理装置 100 に送信する。またディスクハートビート信号 232 の複製の検知の結果を情報処理装置 100 に送信する。

【0055】

===活性化／不活性化メッセージ===

次に本実施の形態に係る活性化／不活性化メッセージ 195 を図 6 に示す。

上述したように活性化／不活性化メッセージ 195 は、メインサイトの情報処理装置 100 の入力装置 150 からオペレータによって入力されたデータに基づいて作成される。活性化／不活性化メッセージ 195 の作成は、オペレータからデータの入力がある毎に行われる。つまりオペレータは、メインサイトのコンピュータシステムとリモートサイトのコンピュータシステムとの間で任意の数の種類のハートビート信号 232 が送受信されるようにすることができる。

【0056】

活性化／不活性化メッセージ 195 は、「活性化／不活性化種別」欄、「ディスクハートビート／ノードハートビート種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄を備える。

【0057】

「活性化／不活性化種別」欄は、ハートビート信号 232 の送受信を行うか行わないかを示す。ハートビート信号 232 の送受信を行うようにすることをハートビート信号 232 を活性化するという。ハートビート信号 232 の送受信を行わないようにすることをハートビート信号 232 を不活性化するという。「活性化／不活性化種別」欄に記載される値としては、例えば "ACTIVATE"、又は "DEACTIVATE" とすることができる。"ACTIVATE" の場合はハートビート信号 232 が活性化され、"DEACTIVATE" の場合はハートビート信号 232 が不活性化されることを示す。

【0058】

「ディスクハートビート／ノードハートビート種別」欄は、活性化又は不活性化されるハートビート信号 232 がノードハートビート信号 232 なのかディスクハートビート信号なのかを示す。「ディスクハートビート／ノードハートビート種別」欄に記載される値としては、例えば "DISK HEARTBEAT"、又は "NODE HEARTBEAT" とすることができる。"DISK HEARTBEAT" の場合は、活性化又は不活性化されるハートビート信号 232 はディスクハートビート信号 232 であることを示し、"NODE HEARTBEAT" の場合は、活性化又は不活性化されるハートビート信号 232 はノードハートビート信号 232 であることを示す。

【0059】

「ローカルデバイスアドレス」欄は、ハートビート信号 232 が書き込まれるハートビート PVOL (230) のアドレス (記憶位置) を示す。

「リモートデバイスアドレス」欄は、ハートビート信号 232 が書き込まれるハートビート SVOL (230) のアドレス (記憶位置) を示す。

【0060】

ハートビート制御プログラム 191 は、オペレータにより入力されたデータに基づいて作成した活性化／不活性化メッセージ 195 をリモートサイトの情報処理装置 100 に送信する。活性化／不活性化メッセージ 195 のリモートサイトの情報処理装置 100 への送信は、第 1 のネットワーク 300 を介して行われるようにすることもできるし、第 2 の

ネットワーク 310 を介して行われるようにすることもできる。第 2 のネットワーク 310 を介して行われるようにする場合には、レプリケーションによって行われるようにすることもできる。

【0061】

===ハートビート状態テーブル===

続いてメインサイトの情報処理装置 100、及びリモートサイトの情報処理装置 100 はそれぞれ、活性化／不活性化メッセージ 195 に基づいて、ハートビート状態テーブル 194 を作成、更新する。本実施の形態に係るハートビート状態テーブルを図 7 に示す。

【0062】

ハートビート状態テーブル 194 は、ノードハートビート信号 232 の状態を示す部分と、ディスクハートビート信号 232 の状態を示す部分とを備える。

ノードハートビート信号 232 の状態を示す部分は、「ノードハートビート状態」欄、「P/S 種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄、「デバイス状態」欄を備える。「P/S 種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄、「デバイス状態」欄は、ノードハートビート信号 232 の種類の数毎に一つ以上設けられる。つまり、ノードハートビート信号 232 の種類は、「P/S 種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄の記載の組み合わせにより特定される。

【0063】

「ノードハートビート状態」欄は、メインサイトのコンピュータシステムとリモートサイトのコンピュータシステムとの間で、ノードハートビート信号 232 の送受信が行われているか否かを示す。「ノードハートビート状態」欄に記載される値としては、「ENABLE」、又は「FAILED」とすることができる。「ENABLE」の場合は、少なくとも一種類以上のノードハートビート信号 232 の送受信が、メインサイトのコンピュータシステムとリモートサイトのコンピュータシステムとの間で行われていることを示す。「FAILED」の場合は、ノードハートビート信号 232 の送受信が全く行われていないことを示す。

【0064】

「P/S 種別」欄は、ノードハートビート信号 232 が送信されるべきものなのか、受信されるべきものなのかを示す。すなわち、引き続き「ローカルデバイスアドレス」欄がハートビート PVOL (230) であるのか、ハートビート SVOL (230) であるのかを示す。「P/S 種別」欄に記載される値としては、「PVOL」、又は「SVOL」とすることができる。活性化／不活性化メッセージ 195 を作成した情報処理装置 100 は「P/S 種別」欄を「PVOL」とする。活性化／不活性化メッセージ 195 を送信された情報処理装置 100 は「P/S 種別」欄を「SVOL」とする。

【0065】

「ローカルデバイスアドレス」欄は、同一コンピュータシステム内のストレージ装置 200 が備える記憶ボリューム 230 における、ノードハートビート信号 232 が書き込まれるアドレスを示す。「P/S 種別」欄が「PVOL」の場合には、「ローカルデバイスアドレス」欄に記載されたアドレスに書き込まれたノードハートビート信号 232 が、レプリケーション制御プログラム 217 によって他のストレージ装置 200 に送信される。「P/S 種別」欄が「SVOL」の場合には、「ローカルデバイスアドレス」欄に記載されたアドレスに、レプリケーション制御プログラム 217 によって他のストレージ装置 200 から送信されたノードハートビート信号 232 が書き込まれる。

【0066】

「リモートデバイスアドレス」欄は、異なるコンピュータシステム内のストレージ装置 200 が備える記憶ボリューム 230 における、ノードハートビート信号 232 が書き込まれるアドレスを示す。

【0067】

「デバイス状態」欄は、ノードハートビート信号 232 が正しく送受信されているか否

かを示す。「デバイス状態」欄に記載される値としては、“ENABLE”、又は“FAILED”とすることができる。“ENABLE”の場合は、ノードハートビート信号232が正しく送受信されていることを示す。“FAILED”の場合は、ノードハートビート信号232が正しく送受信されていないことを示す。

【0068】

全ての種類のノードハートビート信号232について、「デバイス状態」欄が“FAILED”となった場合は、ノードハートビート信号232の全ての送受信が正しく行われていないことになるので、「ノードハートビート状態」欄には“FAILED”が記載される。

なお、ノードハートビート信号232が正しく送受信されているか否かについての判断方法については後述する。

【0069】

ディスクハートビート信号232の状態を示す部分は、「ディスクハートビート状態」欄、「P/S種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄、「デバイス状態」欄を備える。「P/S種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄、「デバイス状態」欄は、ディスクハートビート信号232の種類の数毎に一つ以上設けられる。つまり、ディスクハートビート信号232の種類は、「P/S種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄の記載の組み合わせにより特定される。

【0070】

「ディスクハートビート状態」欄は、メインサイトのコンピュータシステムとリモートサイトのコンピュータシステムとの間で、ディスクハートビート信号232の送受信が行われているか否かを示す。「ディスクハートビート状態」欄に記載される値としては、“ENABLE”、又は“FAILED”とすることができる。“ENABLE”の場合は、少なくとも一種類以上のディスクハートビート信号232の送受信が、メインサイトのコンピュータシステムとリモートサイトのコンピュータシステムとの間で行われていることを示す。“FAILED”の場合は、ディスクハートビート信号232の送受信が全く行われていないことを示す。

【0071】

「P/S種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄、「デバイス状態」欄については、ノードハートビート信号232の状態を示す部分と同様である。

【0072】

なお、マイクロ制御VOL230に記憶されるハートビート状態テーブル194は、ディスクハートビート信号232の状態を示す部分のみとすることもできる。これにより、ディスクハートビート信号232の状態をより少ない記憶容量で管理することが可能となる。

【0073】

===ハートビート信号===

ノードハートビート制御プログラム192は、メモリ120に記憶されたハートビート状態テーブル194に基づいて、ノードハートビート信号232を作成し、ストレージ装置200に対して、所定時間以内の間隔、例えば1分以内の間隔で、繰り返し、ハートビートPVOL(230)に対するノードハートビート信号232の書き込み要求を送信する。ノードハートビート信号232は、ストレージ装置200で実行されるストレージ制御プログラム218によってハートビートPVOL(230)に書き込まれる。

【0074】

またディスクハートビート制御プログラム215は、マイクロ制御VOL(230)に記憶されたハートビート状態テーブル194に基づいて、ディスクハートビート信号232を作成し、所定時間以内の間隔、例えば1分以内の間隔で、繰り返し、ハートビートPVOL(230)に対してディスクハートビート信号232の書き込みを行う。

【0075】

本実施の形態に係るノードハートビート信号232、及びディスクハートビート信号232を図8に示す。図8に示すように、本実施の形態に係るノードハートビート信号232、及びディスクハートビート信号232は、「通し番号」欄、「時刻」欄、「ディスクハートビート／ノードハートビート種別」欄、「ローカルデバイスアドレス」欄、「リモートデバイスアドレス」欄の少なくともいずれかを含んで構成される。

【0076】

「通し番号」欄は、ハートビートPVOL(230)に所定時間以内の間隔で繰り返し書き込まれる同一種類のハートビート信号232のそれぞれを識別するための識別情報が記載される欄である。通し番号は、例えばハートビートPVOL(230)に順に書き込まれるハートビート信号232毎に1ずつ値が増えてゆくように付される番号とすることができる。

【0077】

「時刻」欄は、ノードハートビート制御プログラム192、あるいはディスクハートビート制御プログラム215によって、ハートビート信号232が作成された時刻を示す情報が記載される欄である。時刻は、例えばハートビート信号232の作成日時とすることもできるし、ある基準日時からの差分を示す情報とすることもできる。

【0078】

「ディスクハートビート／ノードハートビート種別」欄は、ノードハートビート制御プログラム192によって作成されたノードハートビート信号232であるのか、ディスクハートビート制御プログラム215によって作成されたディスクハートビート信号232であるのかを示す情報が記載される欄である。

【0079】

「ローカルデバイスアドレス」欄は、当該ハートビート信号232が書き込まれるハートビートPVOL(230)の記憶位置を示す情報が記載される欄である。

「リモートデバイスアドレス」欄は、当該ハートビート信号232が書き込まれるハートビートSVOL(230)の記憶位置を示す情報が記載される欄である。

【0080】

===ノードハートビート信号が正しく送受信されているか否かの確認===

リモートサイトの情報処理装置100で実行されるノードハートビート制御プログラム192は、ハートビート状態テーブル194を参照し、「P/S種別」欄が”SVOL”であり、「デバイス状態」欄が”ENABLE”であるものについて、それぞれ、「ローカルデバイスアドレス」欄に記載されているアドレスを指定して、ストレージ装置200に対して、ノードハートビート信号232の複製の読み出し要求を送信する。この読み出し要求の送信は、所定時間毎に、例えば1分間隔で行うようにすることができる。このノードハートビート信号232の複製は、ストレージ装置200によってハートビートSVOL(230)から読みだされ、情報処理装置100に送信される。そしてノードハートビート制御プログラム192は、ハートビートSVOL(230)から読みだされたノードハートビート信号232の複製の内容と、ハートビート状態テーブル194におけるそのノードハートビート信号232に対応する内容とを比較することにより、ノードハートビート信号232が正しく送受信されているか否かを判断する。

【0081】

具体的には、ノードハートビート信号232の「リモートデバイスアドレス」欄の記載内容と、ハートビート状態テーブル194におけるそのノードハートビート信号232に対応する「リモートデバイスアドレス」欄の記載内容が一致するか否か、ノードハートビート信号232の「ローカルデバイスアドレス」欄の記載内容と、ハートビート状態テーブル194におけるそのノードハートビート信号232に対応する「ローカルデバイスアドレス」欄の記載内容が一致するか否か、ノードハートビート信号232の「ディスクハートビート／ノードハートビート種別」欄の記載内容が”NODE HEARTBEAT”であるか否かを比較する。さらに、ノードハートビート信号232の「通し番号」欄に

記載された値が、前回に受信したノードハートビート信号 232 よりも増加しているか否かを確認する。

【0082】

ノードハートビート信号 232 とハートビート状態テーブル 194 の上記各欄の記載内容が全て一致し、かつ、ノードハートビート信号 232 の「ディスクハートビート／ノードハートビート種別」欄の記載内容が” NODE HEARTBEAT” であり、かつ、ノードハートビート信号 232 の「通し番号」欄に記載された値が、前回に受信したノードハートビート信号 232 よりも増加している場合に、ノードハートビート制御プログラム 192 は、ノードハートビート信号 232 が正しく送受信されていると判断する。

そしてノードハートビート制御プログラム 192 は、ノードハートビート信号 232 が正しく送受信されているか否かに応じて、ハートビート状態テーブル 194 の「デバイス状態」欄に” ENABLE”、あるいは” FAILED” を記載する。

ノードハートビート制御プログラム 192 は、ハートビート状態テーブル 194 のノードハートビート信号 232 の状態を示す部分について、全ての「デバイス状態」欄が” FAILED” である場合には、「ノードハートビート状態」欄に” FAILED” を記載する。

【0083】

=== ディスクハートビート信号が正しく送受信されているか否かの確認 ===

ディスクハートビート制御プログラム 215 は、マイクロ制御 VOL (230) に記憶されたハートビート状態テーブル 194 を参照し、「P/S 種別」欄が” SVOL” であり、「デバイス状態」欄が” ENABLE” であるものについて、それぞれ、「ローカルデバイスアドレス」欄に記載されているハートビート SVOL (230) のアドレスから、ディスクハートビート信号 232 の複製を読み出す。この読み出しは、所定時間毎に、例えば 1 分間隔で行うようにすることができる。そしてディスクハートビート制御プログラム 215 は、ハートビート SVOL (230) から読みだされたディスクハートビート信号 232 の複製の内容と、ハートビート状態テーブル 194 におけるそのディスクハートビート信号 232 に対応する内容とを比較することにより、ディスクハートビート信号 232 が正しく送受信されているか否かを判断する。

【0084】

具体的には、ディスクハートビート信号 232 の「リモートデバイスアドレス」欄の記載内容と、ハートビート状態テーブル 194 におけるそのディスクハートビート信号 232 に対応する「リモートデバイスアドレス」欄の記載内容が一致するか否か、ディスクハートビート信号 232 の「ローカルデバイスアドレス」欄の記載内容と、ハートビート状態テーブル 194 におけるそのディスクハートビート信号 232 に対応する「ローカルデバイスアドレス」欄の記載内容が一致するか否か、ディスクハートビート信号 232 の「ディスクハートビート／ノードハートビート種別」欄の記載内容が” DISK HEARTBEAT” であるか否かを比較する。さらに、ディスクハートビート信号 232 の「通し番号」欄に記載された値が、前回に受信したディスクハートビート信号 232 よりも増加しているか否かを確認する。

【0085】

ディスクハートビート信号 232 とハートビート状態テーブル 194 の上記各欄の記載内容が全て一致し、かつ、ディスクハートビート信号 232 の「ディスクハートビート／ノードハートビート種別」欄の記載内容が” DISK HEARTBEAT” であり、かつ、ディスクハートビート信号 232 の「通し番号」欄に記載された値が、前回に受信したディスクハートビート信号 232 よりも増加している場合に、ディスクハートビート制御プログラム 192 は、ディスクハートビート信号 232 が正しく送受信されていると判断する。

そしてディスクハートビート制御プログラム 215 は、ディスクハートビート信号 232 が正しく送受信されているか否かに応じて、ハートビート状態テーブル 194 の「デバイス状態」欄に” ENABLE”、あるいは” FAILED” を記載する。

ディスクハートビート制御プログラム 215 は、ハートビート状態テーブル 194 のディスクハートビート信号 232 の状態を示す部分について、全ての「デバイス状態」欄を "FAILED" である場合には、「ディスクハートビート状態」欄に "FAILED" を記載する。

【0086】

===メインサイトのコンピュータシステムの動作状態の判定===

リモートサイトの情報処理装置 100 で実行されるハートビート制御プログラム 191 は、リモートサイトのストレージ装置 200 に対して、マイクロ制御 VOL (230) に記憶されたハートビート状態テーブル 194 の読み出し要求を送信する。この読み出し要求の送信は、所定時間毎に、例えば 1 分間隔で行うようにすることができる。そしてストレージ装置 200 は、マイクロ制御 VOL (230) に記憶されたハートビート状態テーブル 194 を情報処理装置 100 に送信する。ハートビート制御プログラム 191 は、ストレージ装置 200 から送信されたハートビート状態テーブル 194 のうち、ディスクハートビート信号 232 の状態を示す部分を、メモリ 120 に記憶されるハートビート状態テーブル 194 に反映する。そしてハートビート制御プログラム 191 は、ノードハートビート信号 232 及びディスクハートビート信号 232 が、メインサイトのコンピュータシステムとリモートサイトのコンピュータシステムとの間で、正しく送受信されているか否かを確認する。これにより、ハートビート制御プログラム 191 は、ディスクハートビート信号 232 の複製の検知の結果と、ノードハートビート信号 232 の複製の検知の結果とに応じて、メインサイトのコンピュータシステムの動作状態を判定する。

【0087】

さらにリモートサイトの情報処理装置 100 は、メインサイトのコンピュータシステムの動作状態に応じてフェイルオーバを行う必要があるか否かを判定し、必要に応じてリモートサイトのコンピュータシステムに、メインサイトのコンピュータシステムにおける情報処理をフェイルオーバさせる。フェイルオーバの実行はクラスタリング制御プログラム 196 により行うことができる。詳細は後述する。

【0088】

===ハートビート信号の活性化／不活性化===

次に、ハートビート信号 232 を活性化する場合の処理の流れ、及びハートビート信号 232 を不活性化する場合の処理の流れについて説明する。まずハートビート信号 232 を活性化する場合の処理の流れについて、図 9 に示すフローチャートを参照しながら説明する。

【0089】

まずハートビート制御プログラム 191 は、メインサイトの情報処理装置 100 の入力装置 150 に対してオペレータから入力されたデータに基づいて、活性化／不活性化メッセージ 195 を作成する。

オペレータからの指示がハートビート信号 232 の活性化である場合には、ハートビート制御プログラム 191 は、ハートビート用ミラーが生成されているか否かを確認する (S1000)。ハートビート用ミラーとは、レプリケーションのペアが形成されているハートビート P VOL (230) とハートビート S VOL (230) とのことをいう。ハートビート用ミラーが生成されていない場合には、ハートビート用ミラーを生成する (S1001)。ハートビート用ミラーの生成は、レプリケーション制御プログラム 217 により行われるようにすることもできる。

【0090】

ハートビート用ミラーが生成されたらハートビート制御プログラム 191 は、メインサイトの情報処理装置 100 のメモリ 120 に記憶されているハートビート状態テーブル 194 を更新する (S1002)。例えば、ハートビート状態テーブル 194 の「ノードハートビート状態」欄を "ENABLE" に変更する。なお、図 9 その他の図面において記載されている生産系サイトはメインサイトである。また待機系サイトはリモートサイトである。

【0091】

つづいてハートビート制御プログラム191は、オペレータから活性化の指示を受けたのがディスクハートビート信号232であるのか否かを確認する(S1003)。もしディスクハートビート信号232の活性化の指示を受けた場合には”YES”に進み、マイクロ制御VOL(230)に記憶されているハートビート状態テーブル194を更新する(S1004)。例えば、ハートビート状態テーブル194の「ディスクハートビート状態」欄を”ENABLE”に変更する。

【0092】

そしてハートビート制御プログラム191は、活性化／不活性化メッセージ195をリモートサイトのコンピュータシステムに送信する(S1005)。リモートサイトのコンピュータシステムの情報処理装置200で実行されるハートビート制御プログラム191は、上記と同様の手順に従い、リモートサイトのコンピュータシステム内のハートビート状態テーブル194を更新する(S1006)。

【0093】

次にハートビート信号232を不活性化する場合の処理の流れについて、図10に示すフローチャートを参照しながら説明する。

【0094】

まずハートビート制御プログラム191は、メインサイトの情報処理装置100の入力装置150に対してオペレータから入力されたデータに基づいて、活性化／不活性化メッセージ195を作成する。

【0095】

オペレータからの指示がハートビート信号232の不活性化である場合には、ハートビート制御プログラム191は、メインサイトの情報処理装置100のメモリ120に記憶されているハートビート状態テーブル194を更新する(S2000)。例えば、ハートビート状態テーブル194の「ノードハートビート状態」欄を”FAILED”に変更する。

【0096】

つづいてハートビート制御プログラム191は、オペレータから不活性化の指示を受けたのがディスクハートビート信号232であるのか否かを確認する(S2001)。もしディスクハートビート信号232の不活性化の指示を受けた場合には”YES”に進み、マイクロ制御VOL(230)に記憶されているハートビート状態テーブル194を更新する(S2002)。例えば、ハートビート状態テーブル194の「ディスクハートビート状態」欄を”FAILED”に変更する。

【0097】

そしてハートビート制御プログラム191は、活性化／不活性化メッセージ195をリモートサイトのコンピュータシステムに送信する(S2003)。リモートサイトのコンピュータシステムの情報処理装置200で実行されるハートビート制御プログラム191は、上記と同様の手順に従い、リモートサイトのコンピュータシステム内のハートビート状態テーブル194を更新する(S2004)。

【0098】

そしてハートビート制御プログラム191は、不要なハートビート用ミラーが存在するか否かを確認する(S2005)。不要なハートビート用ミラーが存在する場合には削除する(S2006)。不要なハートビート用ミラーの削除は、レプリケーション制御プログラム217により行われるようにすることもできる。

【0099】

===ノードハートビート信号の送信===

以上の処理により、ノードハートビート信号232の活性化が行われると、メインサイトのコンピュータシステムからリモートサイトのコンピュータシステムに対するノードハートビート信号232の送信が開始される。ノードハートビート信号232を送信する際の処理の流れについて、図11に示すフローチャートを参照しながら説明する。

【0100】

まずノードハートビート制御プログラム192は、メモリ120に記憶されたハートビート状態テーブル194のうち、ノードハートビート信号232の状態を示す部分を参照し、「P/S種別」欄が”PVOL”であり、かつ、「デバイス状態」欄が”ENABLE”であるものについて、「ローカルデバイスアドレス」欄に記載されたハートビートPVOL(230)のアドレスを特定する(S3000)。そしてそのハートビートPVOL(230)のアドレスを指定して、ストレージ装置200に対して、ノードハートビート信号232の書き込み要求を送信する。このノードハートビート信号232は、ストレージ装置200で実行されるストレージ制御プログラム218によってハートビートPVOL(230)に書き込まれる(S3001)。そしてレプリケーション制御プログラム217は、ペア管理テーブル216に基づいて、ハートビートPVOL(230)に書き込まれたノードハートビート信号232を、第2のネットワーク310を介してリモートサイトのストレージ装置200に送信する。レプリケーション制御プログラム217は、第2のネットワーク310を介してメインサイトのストレージ装置200から送信されたノードハートビート信号232を、ハートビートSVOL(230)に書き込む(S3002)。これにより、メインサイトのコンピュータシステムからリモートサイトのコンピュータシステムに対するノードハートビート信号232の送信を行うことができる。なお図11その他の図面において記載されている遠隔ミラーはレプリケーション制御プログラム217である。

【0101】

なお以上の処理は、メモリ120に記憶されたハートビート状態テーブル194のノードハートビート信号232の状態を示す部分の「P/S種別」欄が”PVOL”であり、かつ、「デバイス状態」欄が”ENABLE”であるもののそれぞれについて行われる。

【0102】

===ディスクハートビート信号の送信===

また、ディスクハートビート信号232の活性化が行われると、メインサイトのコンピュータシステムからリモートサイトのコンピュータシステムに対するディスクハートビート信号232の送信が開始される。ディスクハートビート信号232を送信する際の処理の流れについて、図12に示すフローチャートを参照しながら説明する。

【0103】

まずディスクハートビート制御プログラム215は、マイクロ制御VOL(230)に記憶されたハートビート状態テーブル194のうち、ディスクハートビート信号232の状態を示す部分を参照し、「P/S種別」欄が”PVOL”であり、かつ、「デバイス状態」欄が”ENABLE”であるものについて、「ローカルデバイスアドレス」欄に記載されたハートビートPVOL(230)のアドレスを特定する(S4000)。そしてそのハートビートPVOL(230)のアドレスに対してディスクハートビート信号232の書き込みを行う(S4001)。そしてレプリケーション制御プログラム217は、ペア管理テーブル216に基づいて、ハートビートPVOL(230)に書き込まれたディスクハートビート信号232を、第2のネットワーク310を介してリモートサイトのストレージ装置200に送信する。レプリケーション制御プログラム217は、第2のネットワーク310を介してメインサイトのストレージ装置200から送信されたディスクハートビート信号232を、ハートビートSVOL(230)に書き込む(S4002)。これにより、メインサイトのコンピュータシステムからリモートサイトのコンピュータシステムに対するディスクハートビート信号232の送信を行うことができる。

【0104】

なお以上の処理は、マイクロ制御VOL(230)に記憶されたハートビート状態テーブル194のディスクハートビート信号232の状態を示す部分の「P/S種別」欄が”PVOL”であり、かつ、「デバイス状態」欄が”ENABLE”であるもののそれぞれについて行われる。

【0105】

===ディスクハートビート信号の受信===

続いてディスクハートビート信号 232 を受信する際の処理の流れについて、図 13 に示すフローチャートを参照しながら説明する。

【0106】

まず、ディスクハートビート制御プログラム 215 は、上述したように、マイクロ制御 VOL (230) に記憶されたハートビート状態テーブル 194 に記載された内容と、ハートビート SVOL (230) に書き込まれたディスクハートビート信号 232 の内容とを比較することにより、ディスクハートビート信号 232 が正常であるか否かを調べる (S5000)。ディスクハートビート信号 232 が正常でない場合は、S5001 にて "NO" に進み、マイクロ制御 VOL (230) に記憶されているハートビート状態テーブル 194 の内容を更新する (S5002)。更新は、ハートビート状態テーブル 194 の「デバイス状態」欄に "FAILED" を記載することにより行う。

【0107】

=== ノードハートビート信号の受信 ===

次にノードハートビート信号 232 を受信する際の処理の流れについて、図 14 に示すフローチャートを参照しながら説明する。

【0108】

まず、ノードハートビート制御プログラム 192 は、上述したように、メモリ 120 に記憶されたハートビート状態テーブル 194 に基づいて、ストレージ装置 200 に対して、ハートビート SVOL (230) に書き込まれたノードハートビート信号 232 の複製の読み出し要求を送信する。このノードハートビート信号 232 の複製は、ストレージ装置 200 で実行されるストレージ制御プログラム 218 によってハートビート SVOL (230) から読みだされ、情報処理装置 100 に送信される。そしてノードハートビート制御プログラム 192 は、メモリ 120 に記憶されたハートビート状態テーブル 194 に記載された内容と、ノードハートビート信号 232 の内容とを比較することにより、ノードハートビート信号 232 が正常であるか否かを調べる (S6000)。ノードハートビート信号 232 が正常でない場合は、S6001 にて "NO" に進み、メモリ 120 に記憶されているハートビート状態テーブル 194 の内容を更新する (S6002)。更新は、ハートビート状態テーブル 194 の「デバイス状態」欄に "FAILED" を記載することにより行う。

【0109】

一方、ハートビート制御プログラム 191 は、ストレージ装置 200 に対して、マイクロ制御 VOL (230) に記憶されたハートビート状態テーブル 194 の読み出し要求を送信する。ハートビート制御プログラム 191 は、ストレージ装置 200 で実行されるストレージ制御プログラム 218 により送信されたハートビート状態テーブル 194 のうち、ディスクハートビート信号 232 の状態を示す部分を、メモリ 120 に記憶されるハートビート状態テーブル 194 に反映する (S6003)。そしてハートビート制御プログラム 191 は、メインサイトのコンピュータシステムの動作状態を判定し、故障部位を特定する (S6004)。

【0110】

故障部位の特定は、例えば図 15 に示すハートビート判定テーブル 198 を参照することにより行うことができる。図 15 において、ノードハートビート信号 232 が「OK」とは、ハートビート状態テーブル 194 の「ノードハートビート状態」欄が "ENABLE" である状態をいう。またノードハートビート信号 232 が「NG」とは、ハートビート状態テーブル 194 の「ノードハートビート状態」欄が "FAILED" である状態をいう。ディスクハートビート信号 232 についても同様である。

【0111】

そして、ノードハートビート信号 232 及びディスクハートビート信号 232 の送受信がそれぞれ正常に行われたか否かに応じて、ハートビート判定テーブル 198 の「1」、「2」、「3」、「4」のいずれかに分類され、故障部位が特定される。

【0112】

例えば、図 16 に示す構成のストレージシステム 400 において、ノードハートビート信号 232 及びディスクハートビート信号 232 の送受信がおこなわれる場合には、図 17 に示すように故障部位が特定できる。

【0113】

図 17 において、丸印が記載された部位は異常ではないことを示し、黒三角印が記載された部位は異常である可能性があることを示す。例えばノードハートビート信号 232 は正常に送受信されているのだが、ディスクハートビート信号 232 が正しく送受信されていない場合には、図 15 に示すハートビート判定テーブル 198 の「3」に分類されるので、メインサイトのストレージ装置 200 のディスク制御装置 210 が異常であるということが即座に判定できる。また例えばディスクハートビート信号 232 は正常に送受信されているのだが、ノードハートビート信号 232 が正しく送受信されていない場合には、図 15 に示すハートビート判定テーブル 198 の「2」に分類されるので、メインサイトの情報処理装置 100 が異常であるか、メインサイトの情報処理装置 100 とストレージ装置 200 とを接続する入出力バス 330 が異常であるかのどちらかであるということが即座に判定できる。

【0114】

また例えば、図 18 に示す構成のストレージシステム 400 においては、図 19 に示すように故障部位を詳細に特定することができる。

【0115】

さらに、例えば図 20 に示すように複数のノードハートビート信号 232 を送受信するようにする場合には、さらに詳細な故障部位の特定が可能となる。この場合のハートビート判定テーブル 198 は図 21 に示すようになる。そして、ノードハートビート信号 232 及びディスクハートビート信号 232 の送受信がそれぞれ正常に行われたか否かに応じて、ハートビート判定テーブル 198 の「1」乃至「8」のいずれかに詳細に分類され、故障部位を詳細に特定することができる。

【0116】

以上のようにして故障部位が特定されると、ハートビート制御プログラム 191 は、メインサイトのコンピュータシステムの動作状態をユーザインタフェースに出力する。また故障部位に応じた処理を実行する (S6005)。動作状態の出力は、例えば図 23 に示すように、情報処理装置 100 の出力装置 160 であるディスプレイの画面上に表示されるウィンドウを通じて行うようにすることができる。また管理端末 260 の出力装置 266 であるディスプレイの画面上に表示されるウィンドウを通じて行うようにすることもできる。このように本実施の形態におけるストレージシステム 400 においては、故障部位が即座に特定され、ユーザインタフェースに出力されるので、故障部位の迅速な特定、保守管理負担の軽減、コンピュータシステムの信頼性の向上を図ることが可能となる。

【0117】

またハートビート制御プログラム 191 は、故障部位に応じて例えばフェイルオーバーを行うか行わないかを決定し、必要な場合にはクラスタリング制御プログラム 196 にフェイルオーバーを行わせるようにすることができる。例えば、図 20 に示すストレージシステム 400 の構成において、ディスクハートビート信号 232 は正常に送受信されているが、2つのノードハートビート信号 232 の送受信がいずれも正常に行われていない場合には、黒三角印の部位に異常が発生しているか否かを、例えば診断プログラムを実行することにより一つずつ確認してゆき、最終的にメインサイトの情報処理装置 200 に異常が発生していると判断した場合には、フェイルオーバーを行うようにすることもできる。

【0118】

以上のように、本実施の形態に係るストレージシステム 400 においては、リモートサイトのコンピュータシステムは、メインサイトのコンピュータシステムからディスクハートビート信号 232 が送信されないことを検知することにより、メインサイトのストレージ装置 200 に異常が発生したことを検知することができる。これにより、メインサイトのストレージ装置 200 の異常をいち早く検知、特定することが可能となる。このため迅

速な障害対応が可能となり、保守管理負担の軽減や、保守サービスの向上、さらにはコンピュータシステムの信頼性の向上を図ることが可能となる。

【0119】

さらに本実施の形態に係るストレージシステム400においては、ディスクハートビート信号232の送受信に加え、ノードハートビート信号232の送受信を組み合わせることにより、コンピュータシステムにおける異常箇所の特特定をより容易に、より詳細に、より正確に特定することが可能となる。またクラスタ制御を行う場合に、障害検出時の動作を特定することが可能となる。例えばハートビート状態テーブル194の「ノードハートビート状態」が”FAILED”であり、「ディスクハートビート状態」が”ENABLE”の場合は、メインサイトの情報処理装置A(100)での情報処理が続行できないと判断し、リモートサイトの情報処理装置B(100)での処理に切り替えることができる。このように、従来は情報処理サービスの提供の継続が不可能となっていた場合にも、情報処理サービスの提供を継続して行うことを可能とすることができる。

【0120】

また本実施の形態に係るストレージシステム400は、これらのハートビート信号232の送受信を高信頼な第2のネットワーク310を介して行うので、信頼性の向上を図ることも可能となるのである。

【0121】

以上発明を実施するための最良の形態について説明したが、上記実施の形態は本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明はその趣旨を逸脱することなく変更、改良され得ると共に、本発明にはその等価物も含まれる。

【図面の簡単な説明】

【0122】

【図1】本実施の形態に係るストレージシステムの全体構成を示すブロック図である。

【図2】本実施の形態に係る情報処理装置の構成を示すブロック図である。

【図3】本実施の形態に係る管理端末の構成を示すブロック図である。

【図4】本実施の形態に係るペア管理テーブルを示す図である。

【図5】本実施の形態に係るストレージシステムの全体構成を示すブロック図である。

【図6】本実施の形態に係る活性化／不活性化メッセージを示す図である。

【図7】本実施の形態に係るハートビート状態テーブルを示す図である。

【図8】本実施の形態に係るハートビート信号を示す図である。

【図9】本実施の形態に係るハートビート信号の活性化の処理の流れを示すフローチャートである。

【図10】本実施の形態に係るハートビート信号の不活性化の処理の流れを示すフローチャートである。

【図11】本実施の形態に係るノードハートビート信号の送信の処理の流れを示すフローチャートである。

【図12】本実施の形態に係るディスクハートビート信号の送信の処理の流れを示すフローチャートである。

【図13】本実施の形態に係るディスクハートビート信号の受信の処理の流れを示すフローチャートである。

【図14】本実施の形態に係るノードハートビート信号の受信の処理の流れを示すフローチャートである。

【図15】本実施の形態に係るハートビート判定テーブルを示す図である。

【図16】本実施の形態に係るストレージシステムの全体構成を示すブロック図である。

【図17】本実施の形態に係るハートビート判定テーブルを示す図である。



【図 18】本実施の形態に係るストレージシステムの全体構成を示すブロック図である。

【図 19】本実施の形態に係るハートビート判定テーブルを示す図である。

【図 20】本実施の形態に係るストレージシステムの全体構成を示すブロック図である。

【図 21】本実施の形態に係るハートビート判定テーブルを示す図である。

【図 22】本実施の形態に係るハートビート判定テーブルを示す図である。

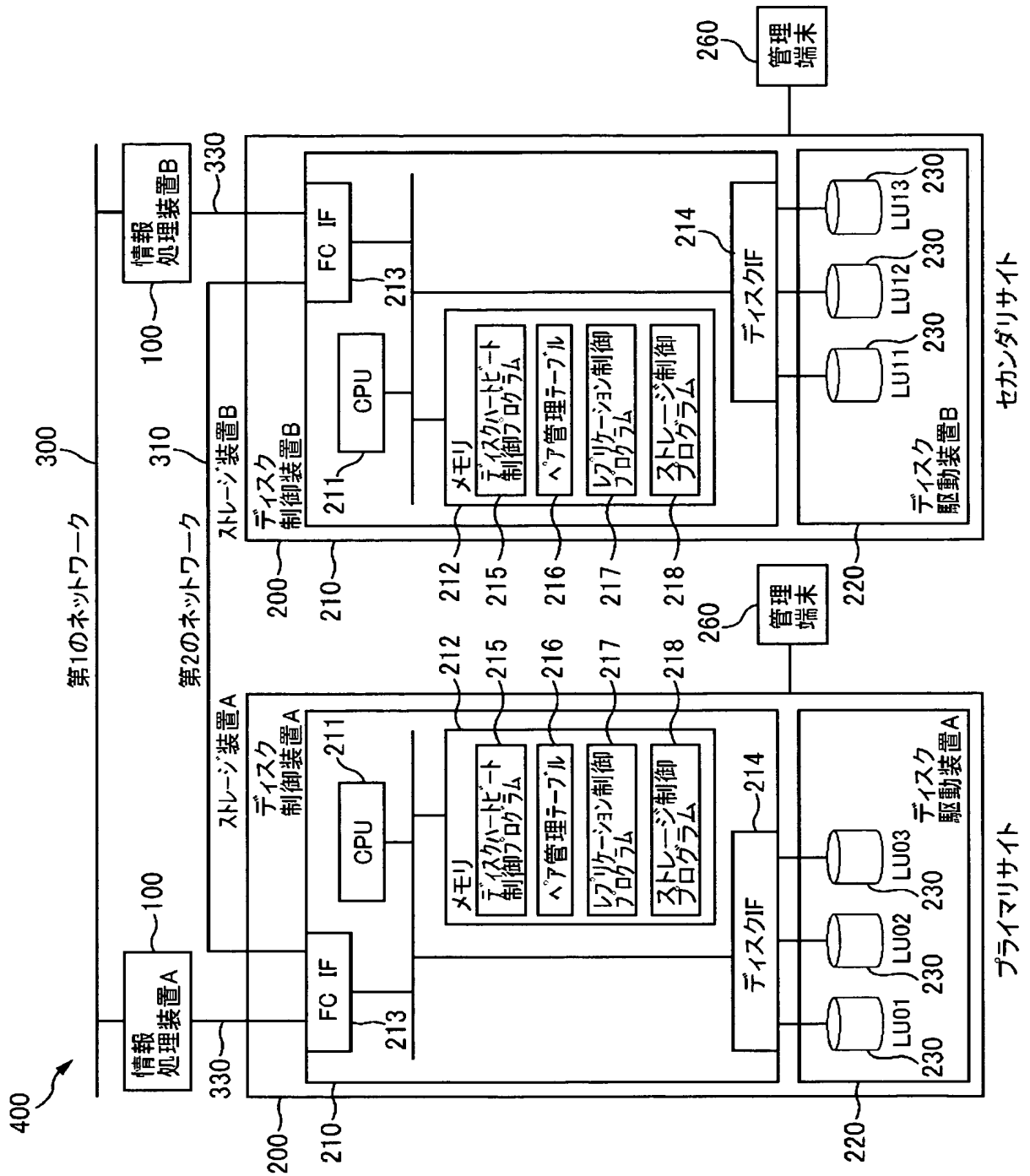
【図 23】本実施の形態に係るコンピュータシステムの動作状態がユーザインタフェースに出力された様子を示す図である。

【符号の説明】

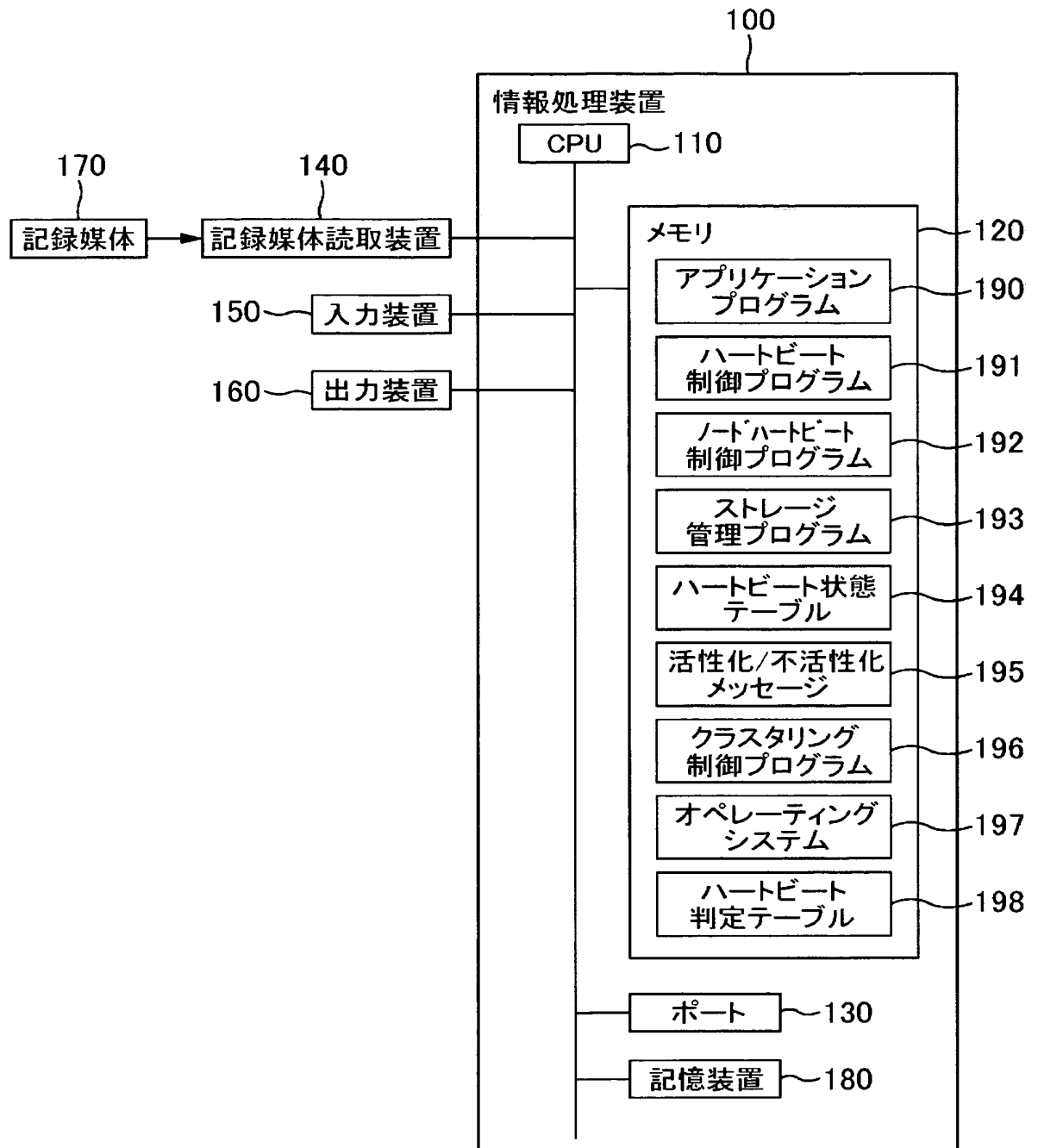
【0123】

100	情報処理装置
191	ハートビート制御プログラム
192	ノードハートビート制御プログラム
193	ストレージ管理プログラム
194	ハートビート状態テーブル
195	活性化不活性化メッセージ
196	クラスタリング制御プログラム
197	オペレーティングシステム
198	ハートビート判定テーブル
200	ストレージ装置
215	ディスクハートビート制御プログラム
216	ペア管理テーブル
217	レプリケーション制御プログラム
218	ストレージ制御プログラム
232	ハートビート信号
260	管理端末
269	制御プログラム
300	第1のネットワーク
310	第2のネットワーク
330	入出力パス
400	ストレージシステム

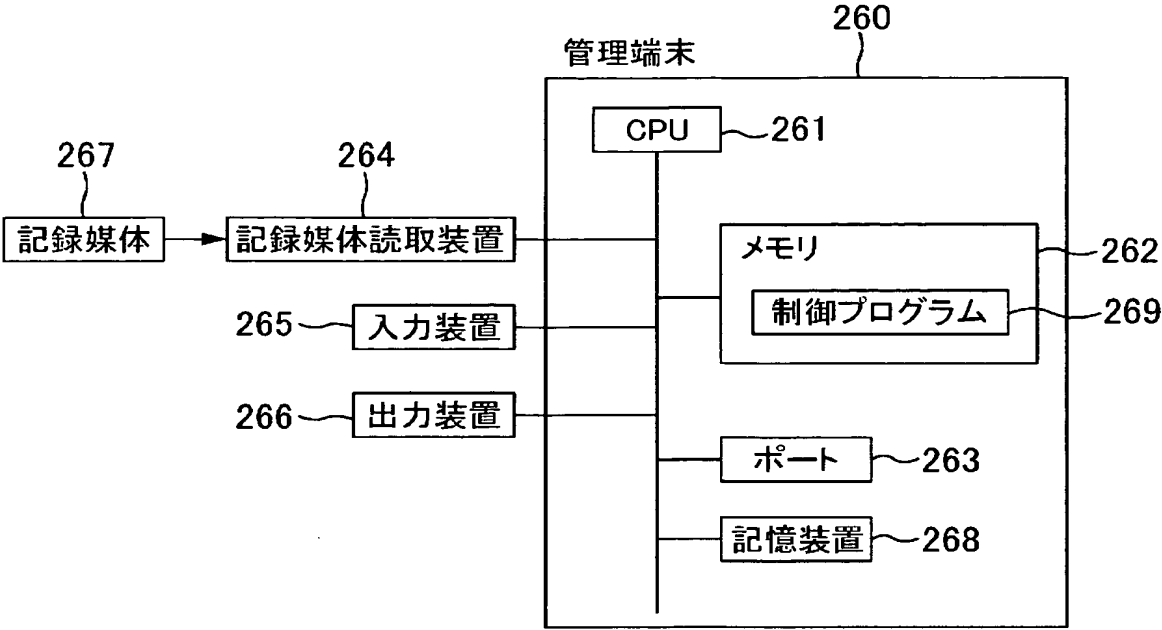
【書類名】 図面
【図 1】



【図 2】



【図 3】



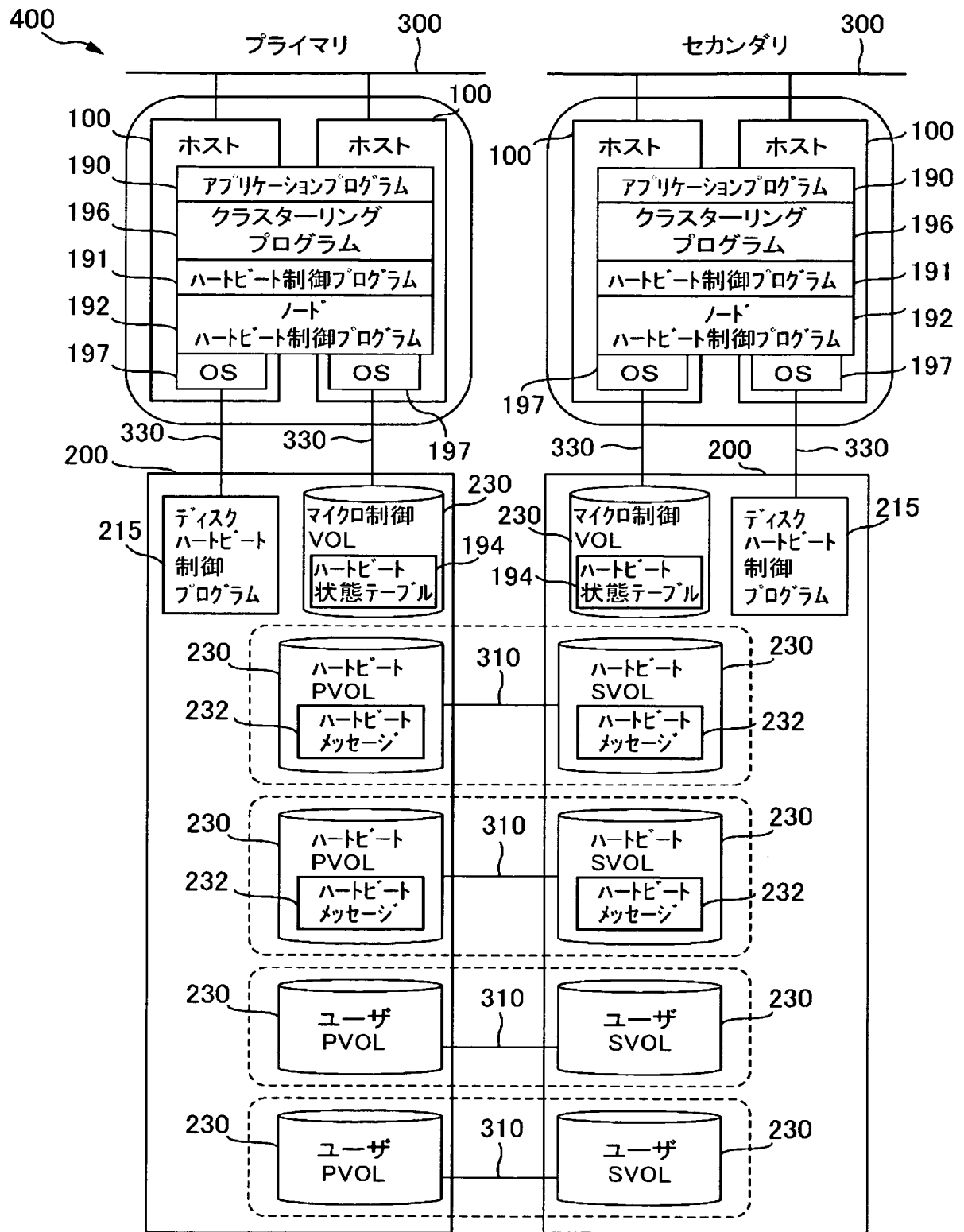
【図 4】

ペア管理テーブル

216

ペア種類	複製方式	複製元装置	複製先装置	複製元ボリューム	複製先ボリューム	ペア状態
リモート	同期	A	B	2	2	スプリット中
リモート	同期	A	B	1	3	ペア中
リモート	同期	A	B	0	0	ペア中
⋮	⋮	⋮	⋮	⋮	⋮	⋮

【図 5】



【図 6】

活性化メッセージ／不活性化メッセージ (195)

活性化／不活性化種別
ディスクハートビート／ノードハートビート
ローカルデバイスアドレス
リモートデバイスアドレス

【図 7】

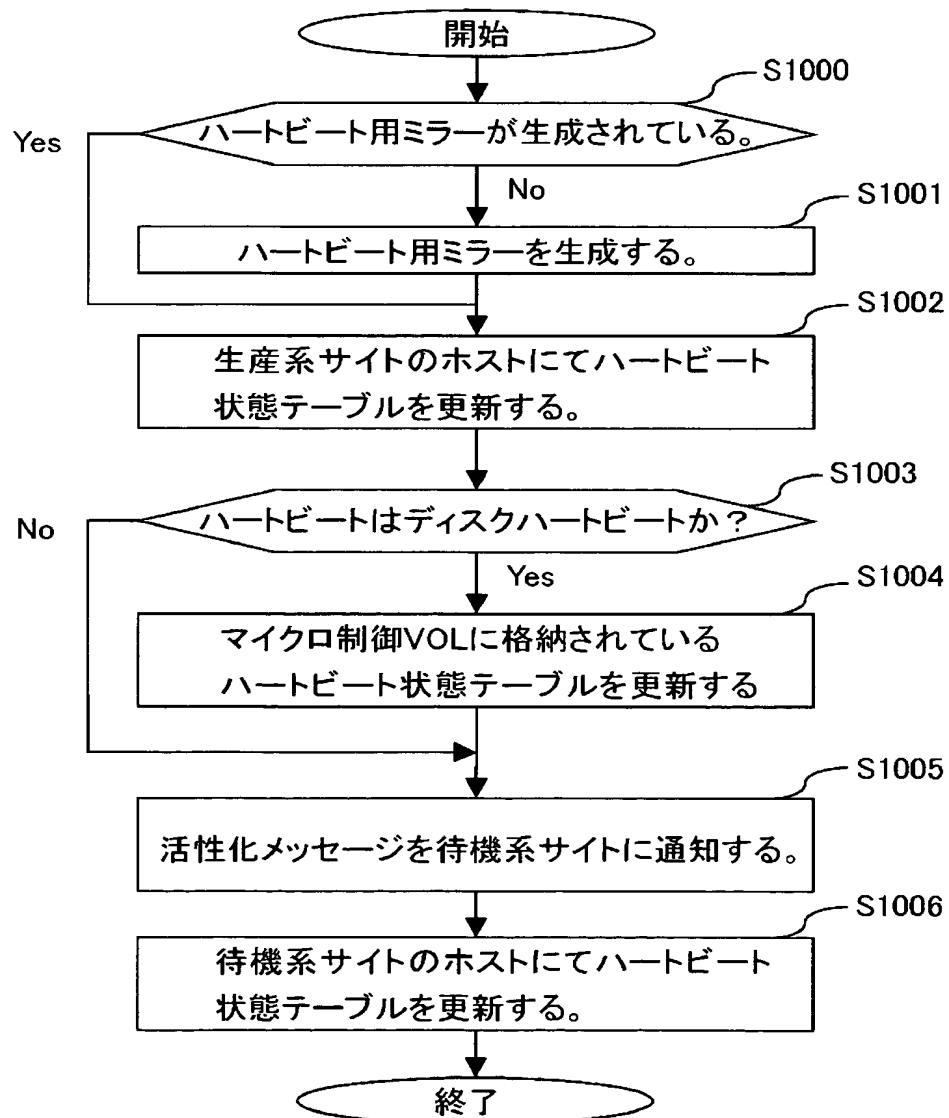
ハートビート状態テーブル(194)

ノードハートビート状態
P/S 種別(1)
ローカルデバイスアドレス (1)
リモートデバイスアドレス (1)
デバイス状態 (1)
P/S 種別(2)
ローカルデバイスアドレス (2)
リモートデバイスアドレス (2)
デバイス状態 (2)
:
ディスクハートビート状態
P/S 種別(3)
ローカルデバイスアドレス (3)
リモートデバイスアドレス (3)
デバイス状態 (3)
P/S 種別(4)
ローカルデバイスアドレス (4)
リモートデバイスアドレス (4)
デバイス状態 (4)
:

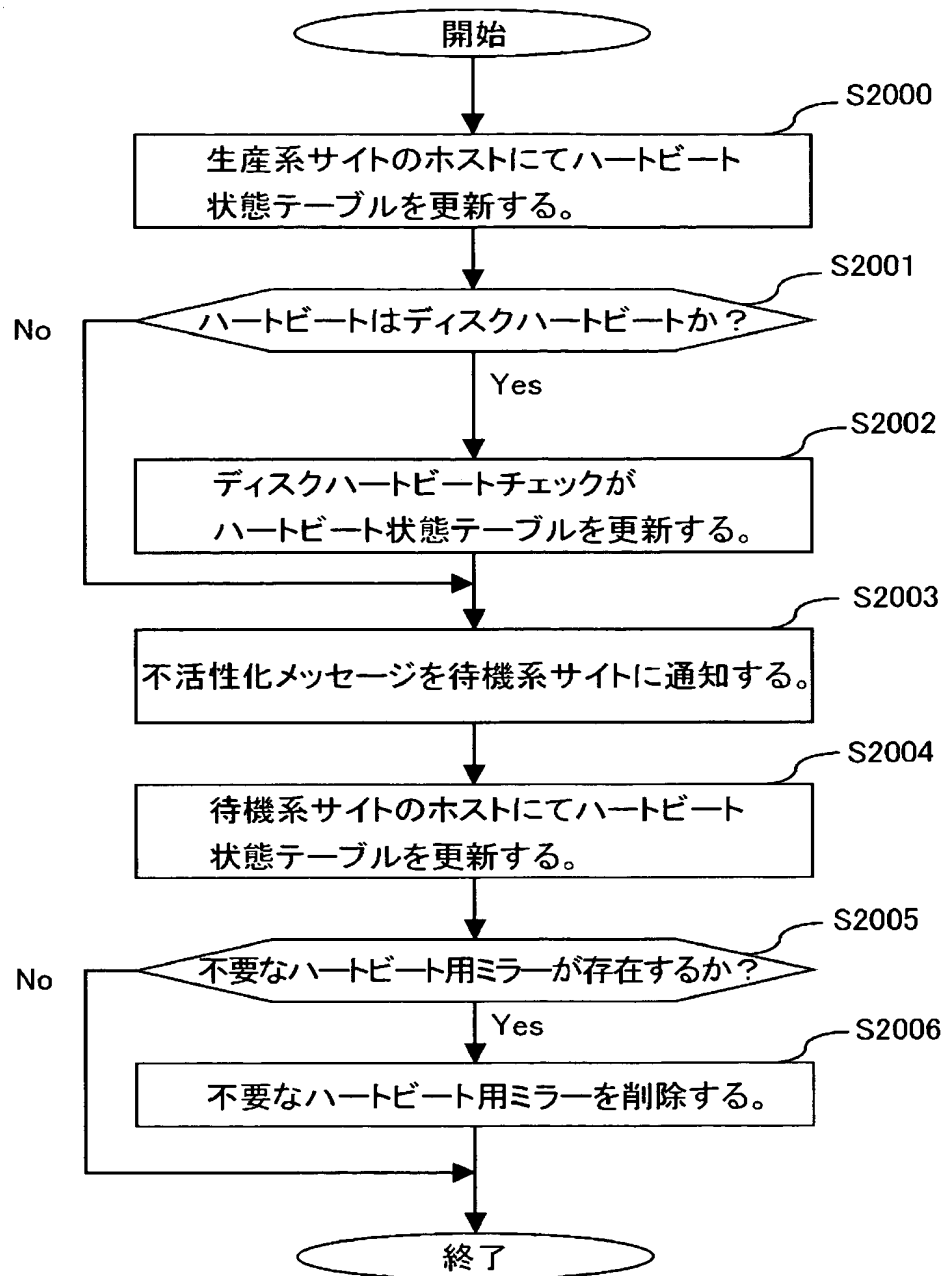
【図 8】

ハートビートメッセージ (232)	
通し番号	
時刻	
ディスクハートビート／ノードハートビート種別	
ローカルデバイスアドレス	
リモートデバイスアドレス	

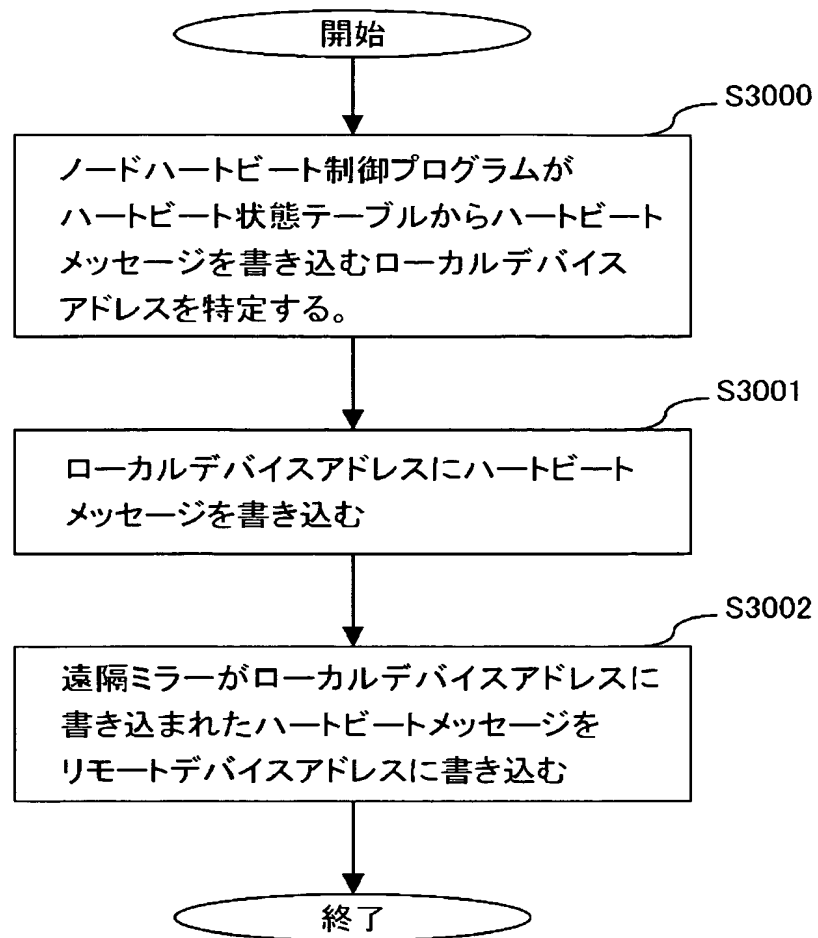
【図 9】



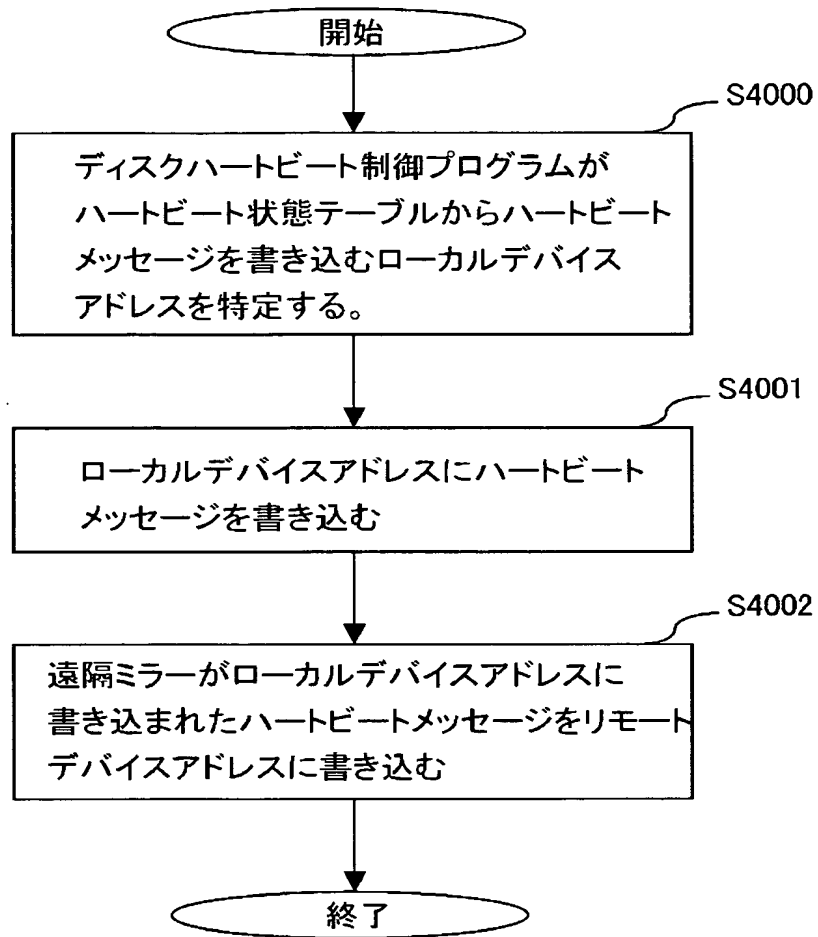
【図 10】



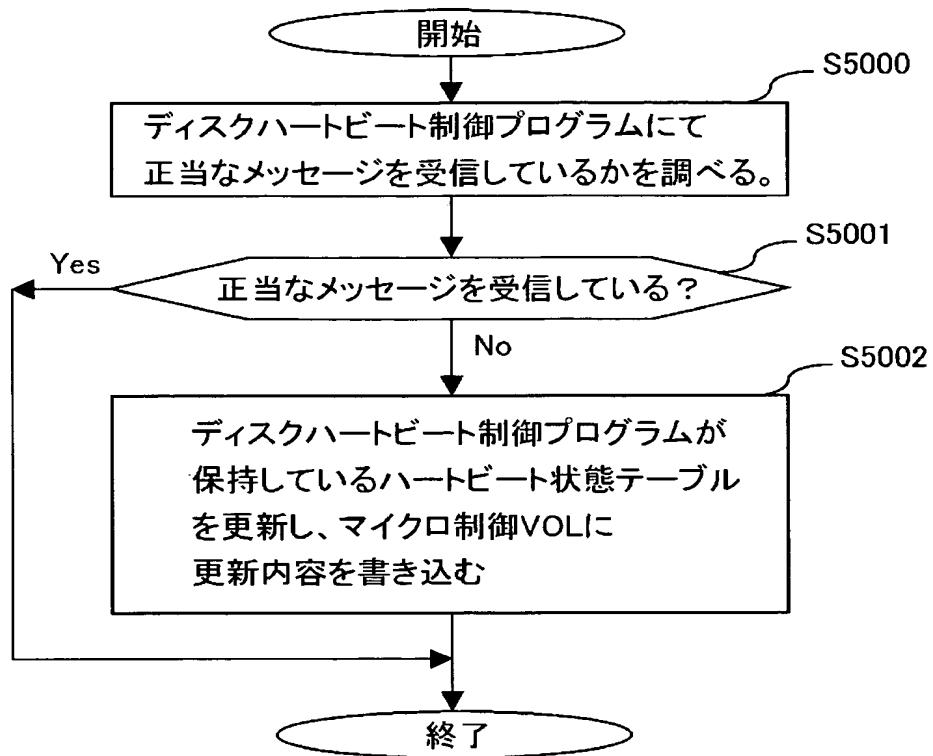
【図 11】



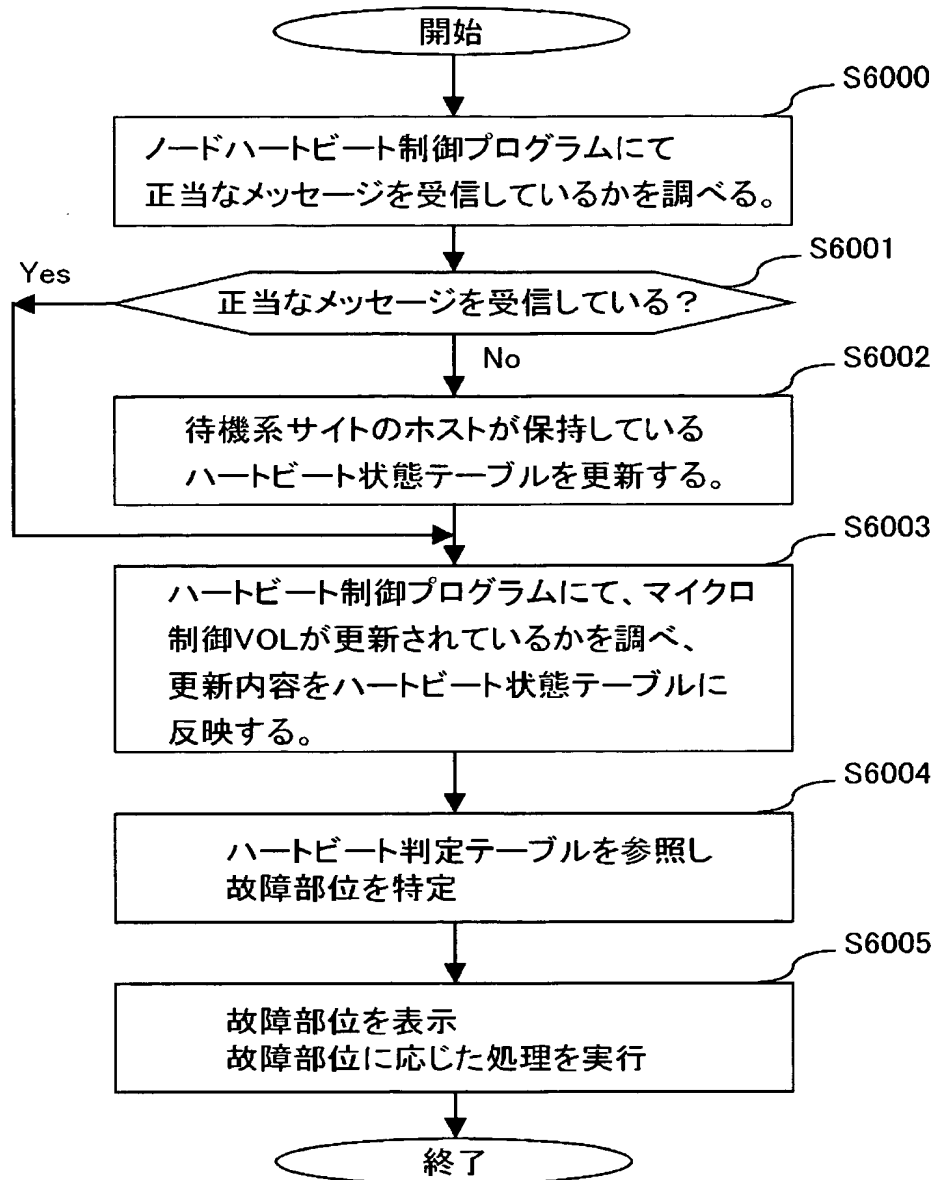
【図 12】



【図 13】



【図 14】

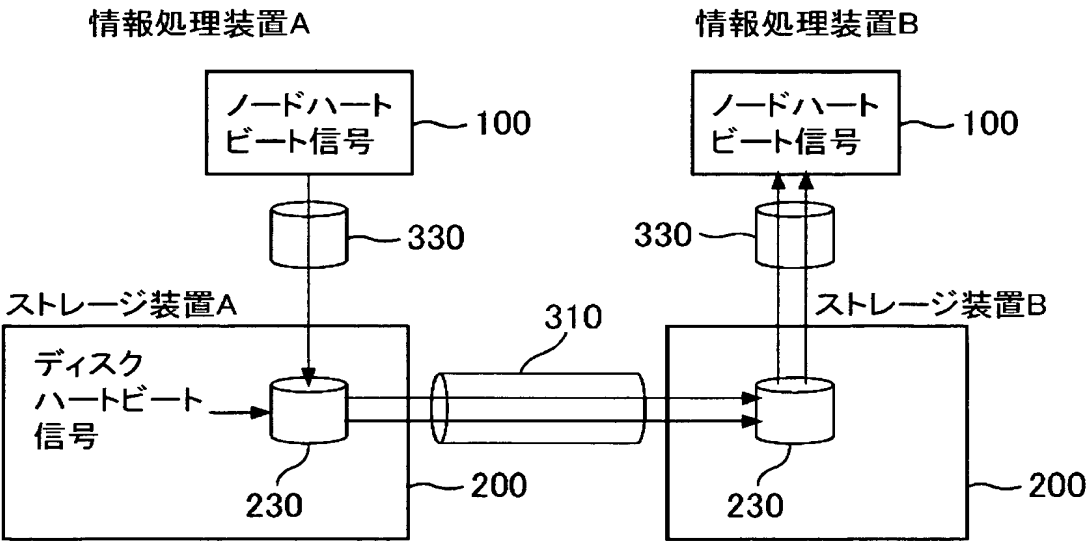


【図 15】

198

		ノードハートビート信号	
		OK	NG
ディスク ハート ビート 信号	OK	①	②
	NG	③	④

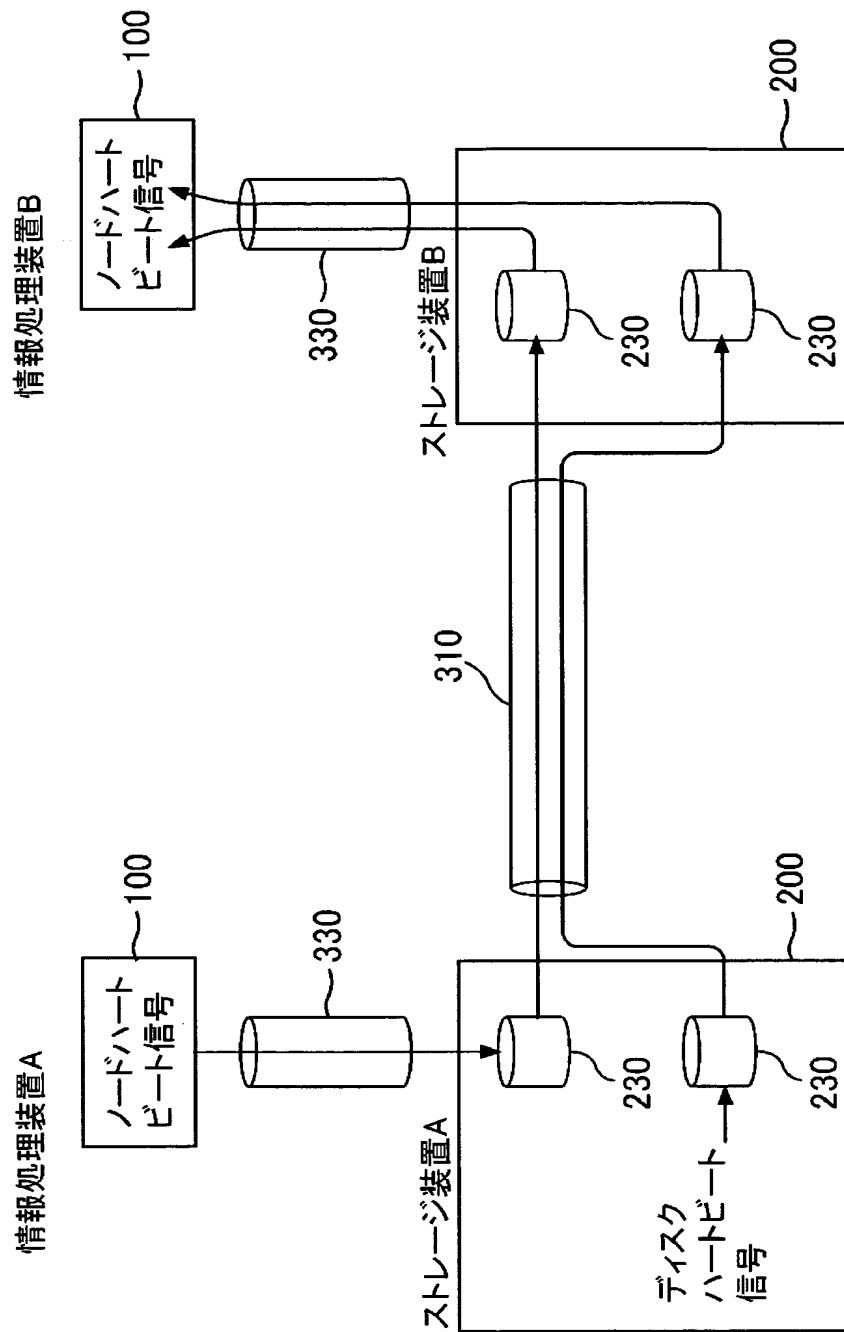
【図 16】



【図 17】

	情報処理 装置A	ストレージ装置A		入出力 パスA	第2の ネットワーク	情報処理 装置B	ストレージ装置B		入出力 パスB
		ディスク制御装置A	P-VOL				ディスク制御装置B	S-VOL	
①	○	○	○	○	○	○	○	○	○
②	▲	○	○	▲	○	○	○	○	○
③	○	▲	○	○	○	○	○	○	○
④	▲	▲	▲	▲	▲	▲	▲	▲	▲

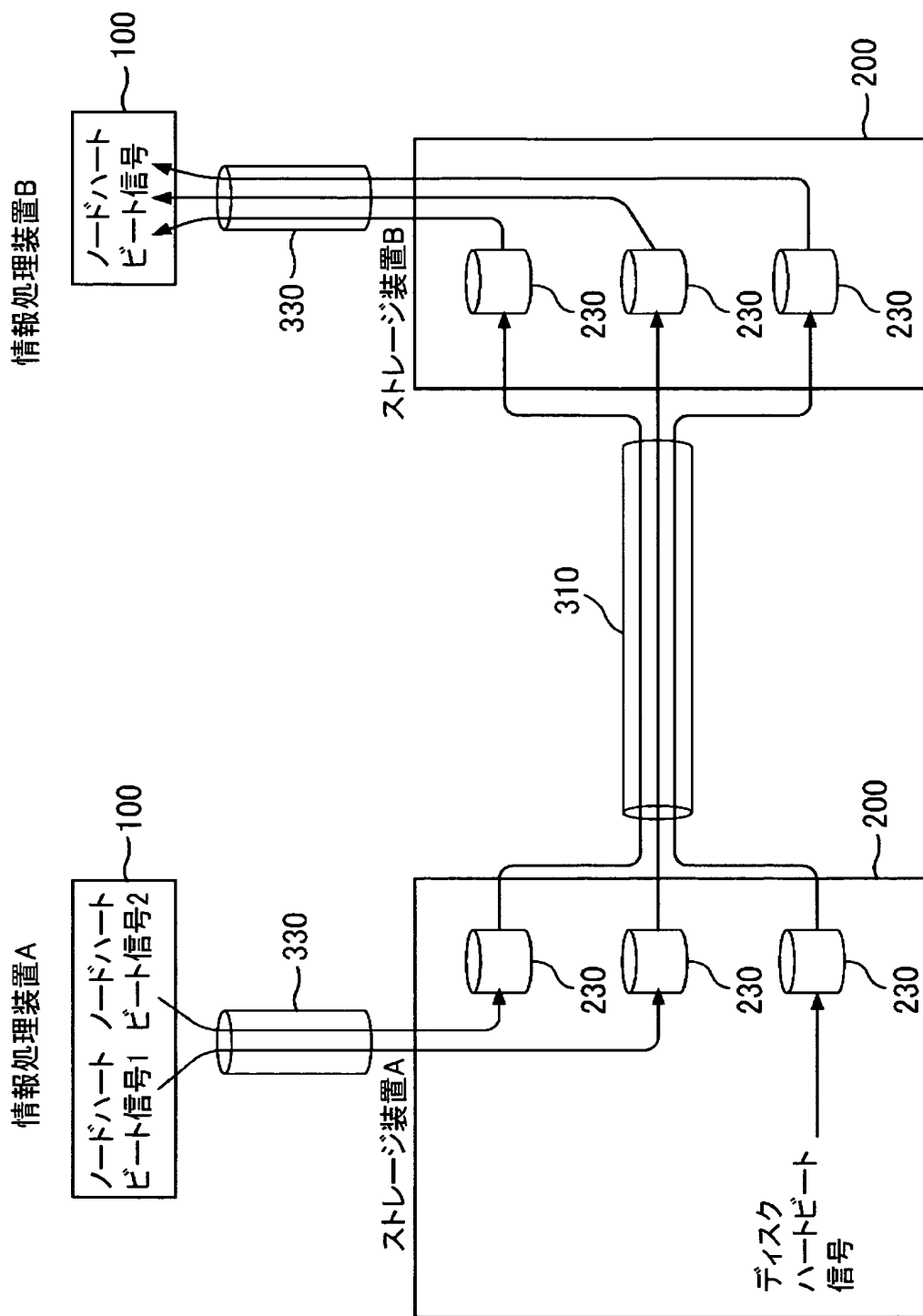
【図 18】



【図 19】

	情報処理 装置A	ストレージ装置A			入出力 パスA	第2の ネットワーク	情報処理 装置B	ストレージ装置B			入出力 パスB
		ディスク 制御装置A	ノードHB用 P-VOL	ディスクHB用 P-VOL				ディスク 制御装置B	ノードHB用 S-VOL	ディスクHB用 S-VOL	
①	○	○	○	○	○	○	○	○	○	○	○
②	▲	○	▲	○	▲	○	○	○	▲	○	○
③	○	▲	○	▲	○	○	○	○	○	▲	○
④	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲

【図 20】



【図 21】

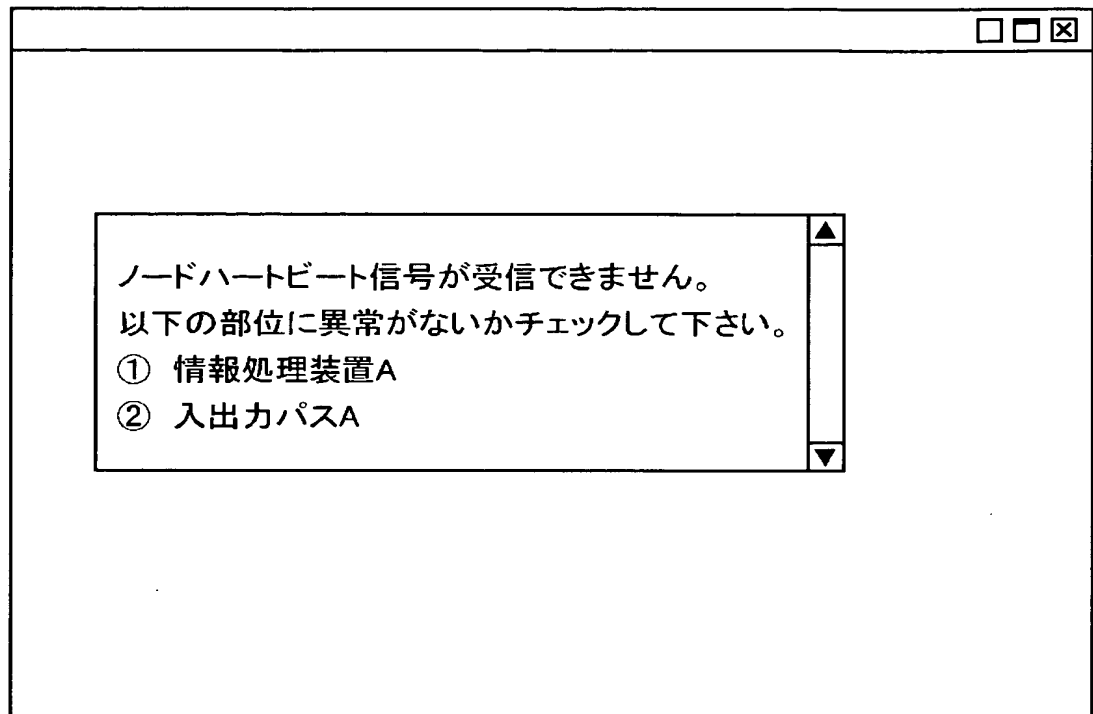
		ノードハートビート信号1			
		OK		NG	
		ノードハートビート信号2		ノードハートビート信号2	
		OK	NG	OK	NG
ディスク ハート ビート 信号	OK	①	②	③	④
	NG	⑤	⑥	⑦	⑧

【図 22】

	情報処理 装置A	ストレージ装置A				入出力 バスA	第2の ネットワーク	情報処理 装置B	ストレージ装置B				入出力 バスB
		ディスク 制御装置A	ノードHB用 P-VOL1	ノードHB用 P-VOL2	ディスクHB用 P-VOL				ディスク 制御装置B	ノードHB用 S-VOL1	ノードHB用 S-VOL2	ディスクHB用 S-VOL	
①	○	○	○	○	○	○	○	○	○	○	○	○	○
②	○	○	○	▲	○	○	○	○	○	○	▲	○	○
③	○	○	▲	○	○	○	○	○	○	▲	○	○	○
④	▲	○	▲	▲	○	▲	○	○	○	▲	▲	○	○
⑤	○	○	○	○	▲	○	○	○	○	○	○	▲	○
⑥	○	○	○	▲	▲	○	○	○	○	○	▲	▲	○
⑦	○	○	▲	○	▲	○	○	○	○	▲	○	▲	○
⑧	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲



【図 23】



【書類名】 要約書**【要約】**

【解決手段】 データを記憶する第1記憶ボリュームを有する第1ストレージ装置と、第1ストレージ装置と通信可能に接続される第2記憶ボリュームを有する第2ストレージ装置とを備え、第1ストレージ装置は第1記憶ボリュームにデータが書き込まれるとその複製を第2ストレージ装置に送信する複製データ送信部を備え、第2ストレージ装置は複製データ送信部により送信されるデータの複製を第2記憶ボリュームに書き込む複製データ受信部を備え、第1ストレージ装置は第1ハートビートメッセージを第1記憶ボリュームに所定時間以内の間隔で繰り返し書き込むディスクハートビート書き込み部を備え、第2ストレージ装置は複製データ受信部により第2記憶ボリュームへ書き込まれる第1ハートビートメッセージの複製の検知を行うディスクハートビート検知部を備えることを特徴とするストレージシステムに関する。

【選択図】 図5



特願 2 0 0 4 - 0 0 2 0 3 7

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所